DOCUMENT RESUME

ED 431 853 UD 033 019

AUTHOR Fiester, Leila
TITLE Getting Smart, Getting Real: Using Research and Evaluation Information To Improve Programs and Policies. Report of the Annie E. Casey Foundation's Research and Evaluation Conference (Baltimore, MD, September 27-29, 1995).
INSTITUTION Annie E. Casey Foundation, Baltimore, MD.
PUB DATE 1995-09-00
NOTE 72p.; "With assistance from Miriam Gonzales and Sara Nathanson."
AVAILABLE FROM Annie E. Casey Foundation, 701 St. Paul Street, Baltimore, MD 21202; Tel: 410-547-6600; Fax: 410-547-6624; Web Site: http://www.aecf.org
PUB TYPE Collected Works - Proceedings (021)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Conferences; Evaluation Methods; *Policy Formation; Program Development; Program Implementation; Research and Development; *Research Utilization; Social Science Research; Urban Problems; *Urban Youth

ABSTRACT

The second research and evaluation conference of the Annie E. Casey Foundation attracted many participants from the public and private sectors. In spite of their diverse backgrounds and approaches to evaluation, the participants shared a sense that researchers and evaluators need to find better ways to collect and use information. The primary goals of the conference was to strengthen the connection between research and the development of programs and policies that improve outcomes for children. This report summarizes the discussions surrounding each of the major research and evaluation issues that conference participants identified as affecting programs and policy making. The conference overview is divided into three sections: (1) understanding the need for better research and evaluation; (2) adapting research and evaluation to meet current needs; and (3) using research and evaluation information to improve programs and policies. Appendixes contain the text of the speeches delivered by the keynote speakers; the conference agenda; and a list of participants. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 431 853

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

WJ Rust

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Getting Smart, Getting Real

Using Research and Evaluation Information to Improve Programs and Policies

Report of the Annie E. Casey Foundation's September 1995 Research and Evaluation Conference

UD 033 019

BEST COPY AVAILABLE

Getting Smart, Getting Real

Using Research and Evaluation Information to Improve Programs and Policies

---

Report of the Annie E. Casey Foundation's September 1995 Research and Evaluation Conference

BEST COPY AVAILABLE

Acknowledgments

Special thanks go to all conference presenters, and particularly, to Donna Schmidt, whose thoroughness made this conference a success. We also wish to thank Leila Fiester of Policy Studies Associates, who authored this report with assistance from Miriam Gonzales and Sara Nathanson.

Additional copies of this report are available from:

The Annie E. Casey Foundation
701 St. Paul Street
Baltimore, MD 21202
410-547-6600
Fax: 410-547-6624
http://www.aecf.org

Research and Evaluation at the Annie E. Casey Foundation

The Annie E. Casey Foundation's "mission, motivation, and message to the world" is its commitment to changing and improving life outcomes for our most disadvantaged children and families, said Tony Cipollone, Associate Director for Education Reform, Research and Evaluation. That belief is rooted in the conviction that outcomes for children will not improve without fundamental, comprehensive, and durable changes in many service and support systems. Current health, education, juvenile justice, and other delivery systems for disadvantaged children and families too often are fragmented, inaccessible, expensive, and irrelevant. They frequently fail to deliver essential services until it is too late, contributing to an overall level of ineffectiveness and to an intergenerational cycle of poverty.

The Foundation operates on the premise that these conditions can be reversed—that "communities can prosper, families can thrive, and children can develop when neighborhoods are supportive, sustaining, and served by systems that are relevant, respectful, and rooted in the communities that they serve," Cipollone said. The Foundation believes that strategic investments in awareness building, capacity development, program demonstrations, and research and evaluation can help move dysfunctional service systems toward greater collaboration, coordination, and flexibility.

In addition to leadership, funding, and other key factors, these changes require accurate, relevant, and compelling information, Cipollone said. Research and evaluation are a conduit for information—and information is power. With information on results provided by evaluation, community stakeholders can make better decisions about organizational practices. Similarly, states informed by research and evaluation can make better decisions about the allocation of resources and policies that affect children and families.

Evaluation plays a major role in the Foundation's theory of change, as a tool for:

- Improving accountability: contributing to understanding about the degree to which interventions represent good judgments about the organizations, communities, and people in which the Foundation places its confidence and resources.
- Revealing the soundness of theories, the practicality of policies, the appropriateness of planning timelines, the relevance of technical assistance, and the extent to which the Foundation has established effective partnerships with grantees.
- Informing funders about the viability of working with states, cities, community-based organizations, and child- and family-serving systems to achieve real transformation and reform.

For these reasons, the Casey Foundation believes that "research and evaluation can, should, and must be a critical and integral component of comprehensive reform strategies."

Table of Contents

Page

Introduction 1

Understanding the Need for Better Research and Evaluation 3

Adapting Research and Evaluation to Meet Current Needs 4

Improving Our Understanding of Outcomes 5

Improving Our Understanding of the Direction of Change 8

Clarifying Attribution of Outcomes 10

Understanding the Reality and Impact of Context 11

Understanding, Acknowledging, and Incorporating Program Participants in Research and Evaluation 13

Using Multiple Strategies to Improve Research and Evaluation 21

Using Research and Evaluation Information to Improve Programs and Policies 22

Understanding How and When Research Influences Policies 22

Producing and Promoting Useful Information 23

Conclusion 28

Appendix A: Text of Speeches A-1

Appendix B: Conference Agenda B-1

Appendix C: Conference Participants C-1

Introduction

The Annie E. Casey Foundation has increasingly linked its efforts to develop comprehensive systems change with a strong research and evaluation agenda. In 1994 the Casey Foundation convened its own evaluators and staff and other researchers to explore key issues in reforming systems and evaluation. The conference focused on making frameworks, technologies, and analysis more responsive to complex, comprehensive efforts to produce change. In 1995, responding to the many questions raised by the first conference, changes in federal priorities, and an evolving national agenda that potentially included extensive changes in services and supports for children and families, the Foundation held a second invitational conference on "Utilizing Research and Evaluation for Programs and Policies."

The second conference, held in Baltimore on September 27-29, attracted an array of participants from the public and private sectors: researchers, program developers and operators, scholars, technical assistance providers, policy makers, evaluators, advocates, representatives of state and federal agencies, and Foundation staff. Despite their diverse backgrounds and approaches to evaluation, the participants shared a sense that researchers and evaluators must find better ways of collecting and using information. As Associate Director for Education Reform, Research and Evaluation Tony Cipollone observed in his opening comments, "The prevailing political rhetoric seems to have been fueled more by perception than fact—by anecdote rather than evidence." In this context, it is especially important to examine how research information can be used to develop effective policies and practices and to foster reforms that better address the needs of children and families.

Goals of the Conference

The primary goal of the conference was to strengthen the connection between research and the ultimate objective of developing programs and policies that improve outcomes for children. The conference provided a forum for:

(1) Sharing new evaluation tools and methods,

(2) Considering policy implications for and of evaluation, and

(3) Discussing four key challenges to evaluating comprehensive services:

- How do we craft useful evaluations in environments that have had negative experiences with research and evaluation? In many situations, the people whose program is being evaluated are afraid to talk to evaluators or do not believe that evaluators have the cultural understanding necessary to evaluate the program.

- How can we help people develop the skills and experiences they need to effectively use the information produced by evaluations? What role does the evaluator play in this process?

- How can we develop better ways to capture short-term, interim benchmarks of change that can be used to see if programs and services are on track?

- What are the best forums and mechanisms to disseminate information gained through research and evaluation so that it reaches a broad audience?

Themes of the Discussion

Keynote speaker Sharon Lynn Kagan, Senior Associate of the Bush Center in Child Development and Social Policy, summed up the overall theme of the conference in her exhortation to researchers and evaluators to "get smart about what doesn't work, get real about new approaches, and get going" on finding new solutions. Several additional themes emerged from panel presentations and topical discussions:

- Researchers and evaluators must develop more collaborative, interactive roles to support better data collection and information sharing.

- Innovative research requires new techniques and approaches—as well as risk-taking in developing these new strategies.

- In order to better use information to improve programs and policies, researchers and evaluators must understand how and when research influences public policy, produce information that is useful to policy makers, and work harder to help policy makers use research and evaluation data.

- The information produced by researchers and evaluators should be presented in simple, compelling ways and targeted to multiple audiences.

This report summarizes the discussions surrounding each of the major research and evaluation issues that conference participants identified as affecting programs and policy making. The overview is divided into three sections: (1) understanding the need for better research and evaluation, (2) adapting research and evaluation to meet current needs, and (3) using research and evaluation information to improve programs and policies. Appendix A contains the text of the speeches delivered by keynote speakers. Appendix B contains the conference agenda. Appendix C contains a list of participants, including their professional affiliations.

Understanding the Need for Better Research and Evaluation

We simply cannot fit the square peg of conventional evaluation into the round hole of comprehensive, community-based efforts.

—Sharon Lynn Kagan

Comprehensive, community-based programs that serve children and families are perched on an "urgent...unparalleled policy precipice," Kagan told conference participants. Programs need measurable results in order to justify their existence and to improve practice. In order to measure these results, however, programs must address the fundamental mismatch between current evaluation design and reform efforts. As many presenters and participants noted:

- Because traditional evaluations were developed to assess narrow, single-issue interventions, policy makers who rely on evaluation data have relied upon standardized, centralized, isolated, and uniform "treatments" rather than the more complex initiatives.

- The community context in which evaluations occur has become increasingly complex. Often, several multi-faceted change efforts are in place, requiring evaluators to use a mixture of strategies and techniques. Conventional evaluation design does not adequately address the complexity of comprehensive services that go beyond limited interventions.

- Random assignment, while integral to traditional research, is not feasible or appropriate for some more innovative programs and initiatives.

In addition, the process of change stimulated by comprehensive programs is difficult to capture:

- The role of independent evaluation in comprehensive, community-based projects is elusive. "We can't get our arms around...what is the treatment, what are the agencies we are evaluating, what are the outcomes and strategies, what is the duration," Kagan observed.

- Change, while desirable for practitioners, complicates research as evaluators struggle to define the scope of an intervention and hold the intervention constant. Further, much of the change process occurs in private, apart from evaluators. And change in one direction may be counteracted by another type of change.

Participants agreed that to resolve these problems and to meet program needs for planning, assessment, and accountability, researchers and evaluators must develop new methods and strategies for evaluating comprehensive community initiatives and programs.

Adapting Research and Evaluation to Meet Current Needs

To meet the need for better research and evaluation, researchers and evaluators must analyze and assess their experiences with various approaches so they can learn from experience, adapt their methods to better fit the changing nature of programs, and share their progress with others. "We need to get to this level quickly if we're going to be risk-takers," observed panel member Heather Weiss. "We're going to have failures. But if we don't take the risks we're not going to get where we want to be collectively."

Innovative research and evaluation include new roles for evaluators and researchers as well as experimentation with new techniques and approaches. These new roles and techniques should focus on six themes, Kagan and other participants suggested: (1) improving our understanding of outcomes; (2) improving our understanding of the direction of change; (3) clarifying attribution of outcomes; (4) understanding the reality and impact of context; (5) understanding, acknowledging, and incorporating program participants in research and evaluation; and (6) using multiple data collection and analysis strategies to make research and evaluation more powerful, comprehensive, and compelling.

Improving Our Understanding of Outcomes

A more exact definition of outcomes would help researchers and evaluators identify program effects. The complexity and flexibility of innovative comprehensive services make it difficult to define outcomes; as some participants noted, even defining programs can be a challenge. Current interventions are multi-dimensional and vary by community. Even specific types of interventions (e.g., family support programs) do not necessarily use uniform, one-dimensional treatments; each program may have different goals, purposes, and interventions. But when evaluators must assess programs that combine many different types of treatments into a complex intervention, the link between specific treatments and outcomes becomes hazy.

To address this challenge, participants proposed several approaches:

- Several participants suggested that evaluators search for indicators that illustrate the domino effect of change, in addition to seeking specific outcomes. For example, participants in a discussion of multi-layered reforms in Missouri suggested examining statewide managed-care reform to see if the new policies reinforce local reforms such as community partnerships.

- Kagan advocated using four broad categories of outcomes. The first assesses the direct impact of the program on children and families and contains information on what children and families know and can do. The second focuses on aggregated information about the conditions that surround children and families; it involves direct observation and statistics collected by service providers. The third category characterizes the services to which children and families have access. The fourth category examines service systems to assess their capacity, infrastructure, and accountability.

- Discussion leader Charles Bruner suggested defining outcomes by examining indicators such as service penetration; family engagement; family growth; community embeddedness; system response, climate for reform, and change; and community-wide family well-being. For example, in one evaluation Bruner identified the "action steps" that families would need to take regarding housing, education, employment, and other categories in order to improve family well-being. Bruner then matched the project's benchmarks of success against progress on these steps.

Clear Outcomes Require Well-Defined Units of Measurement

The challenge of defining outcomes in complex, comprehensive initiatives requires new attention to issues of measurement. Appropriate units of measurement for change are not always clear. For example, should researchers focus on geographic boundaries or on social communities? Traditional methods for defining a unit of measurement—using a geographically defined neighborhood, for example—are not always appropriate in an age when people form communities through the social or ethnic organizations that pull them together, rather than through geographic proximity. If the "neighborhood" is the unit of measure, where exactly do the boundaries of the neighborhood lie? And, as discussion leader and independent researcher Joy Dryfoos noted, what about interventions that target and serve relatively small numbers of children—like many school-based projects? Their effects may easily be "lost" within larger student communities. And evaluators may have trouble distinguishing between users and non-users of school-based or school-linked services.

To resolve these issues, researchers must find different strategies for studying different communities. Often, smaller units of measurement are more effective than large ones. Evaluators may also use different boundaries for measuring different aspects of life (e.g., school, home life, etc.).

An Evaluation of Community Change Considers Residents' Views on Neighborhood Boundaries

Measurement units such as neighborhood boundaries are important to evaluators of community-based initiatives because they influence the evaluator's conception of what constitutes a community. But neighborhood boundaries often are amorphous and hard to articulate. An evaluation led by researcher Claudia Coulton, for example, started with a block group as a tentative proxy for neighborhood and asked residents how they viewed the boundaries of their neighborhood. In some areas, the researchers found a fair amount of consensus; residents even had names for their neighborhoods. In other areas, residents' views of their neighborhood boundaries varied greatly. "You can't say you're going to draw the boundaries where neighbors say they are, because the fact that they can't [agree on] boundaries often means there's something to study there," explained Coulton.

Measurement is especially difficult when evaluators are trying to measure changes in institutions, systems, and communities, rather than simply changes in the behavior of individuals. Yet these broader changes are major goals of community-based, collaborative initiatives, and conference participants agreed that evaluators should find ways to measure them. Several participants noted that evaluators struggle with four issues in particular: (1) defining the type of change, (2) measuring both the means and the end of change (the process and the outcomes), (3) selecting measures of change that capture the quality of the collaboration, and (4) building institutional capacity for self-evaluation. Faced by these challenges, evaluators must continually think in terms of community- and system-level change.

The type of change sought by an initiative also will affect the unit of measurement that evaluators use. For example, people may seek changes in the general population and in the community itself, not solely changes in outcomes for individuals. People may want more solidarity, integration, and civic pride in their community; better exchanges of information among community members; and changes in institutional structures and power relationships. A researcher's methods consequently will vary based on the decision to measure the community itself, a population within the community, or institutional and organizational structures. As Coulton noted:

> If part of what you're after is to have people experience something in the neighborhood that changes them, and you have a highly mobile neighborhood, you may need to track people as they [move] so you don't lose outcome measures. Obviously, we can't measure everything; but if that outcome is your primary focus, you'll have to do some follow-up.

Short-Term or Interim Indicators Provide Valuable Measurements of Progress Toward Goals

Evaluators often face pressures from research funders who want "hard" outcomes and quick answers. These expectations require evaluators to assess a program's achievement before it has had a chance to effect major changes. Most participants agreed that evaluators should identify short-term and interim indicators of progress that test "intermediate hypotheses" in order to assess progress toward key outcomes. "If we're going to try to change circumstances and opportunities, it's not realistic to think we can do it over the short term. We have to take a generational perspective, and we also have to have some markers [of progress] along the way," suggested Bruner.

Lisbeth Schorr suggested in her keynote speech[1] that two types of interim measures can predict later outcomes: "indicators that attach to children, families, and communities and...are a short-term manifestation of long-term outcomes, and indicators of a community's capacity to achieve the identified long-term outcomes." Short-term indicators can take many forms. For example, indicators of early changes in communities may include increased citizen participation, development of networks and relationships among institutions within neighborhoods, the emergence of new leadership, the development of cross-community dialogue, increased decision-making capacity in the community, and a new sense or locus of power. In order to establish a baseline for assessing progress, evaluators should try to measure these indicators early, before they begin to change, Coulton suggested.

Although most participants said interim measures offer useful feedback for practitioners, project administrators, and funders, some were frustrated by the responsibility of defining and collecting short-term indicators while concentrating on long-term outcomes. "Does it make sense to talk about intermediate outcomes when institutional and children and family outcomes [are] the ultimate outcomes? This is very complex, and trying to attribute cause is practically impossible," said one researcher.

Further, as Schorr pointed out, knowledge about the connections between short-term indicators of community capacity and long-term outcomes is "at a more primitive stage" than evaluators' understanding about relationships between interim and long-term indicators for children and families. "One useful next step would be to systematically examine findings in the recent literature and ongoing experience to provide a more rigorous and deeper understanding" of these connections, she suggested.

Improving Our Understanding of the Direction of Change

Becoming clear on outcomes means reaching a better understanding of "pathways of change"—figuring out how to attribute change appropriately, given the multitude of dependent and independent variables. To understand change pathways, evaluators must address many

---

[1]Schorr was one of the scheduled keynote speakers at the conference but was unable to attend. Her written speech was delivered by Anne Kubisch, Director of the Aspen Institute Roundtable on Comprehensive Community Initiatives for Children and Families.

issues: What do they want to measure, how do they expect change to occur, what kinds of change do they expect, and how do they expect changes to be related to one another? "Change in communities may not be linear," cautioned one participant. "It may feed on itself in a reciprocal way."

A focus on theories of change—the goals, beliefs, and expectations that drive programs and policies—gives evaluators new tools for assessing the direction of change. As researcher Carol Weiss has proposed, theory-based evaluation supplements quantitative studies and provides an effective alternative to research based on random experiments, which are often impractical in community-based initiatives. After identifying the theories of change, an evaluator using a theory-based approach works with project staff to identify interim steps that, based on experience and research, link the elements of the theory together. There is no consistent recipe for identifying theories of change; in most cases, researchers must simply talk to program leaders, staff, parents, and/or community leaders to find out what they are doing and what they hope to accomplish in the short term and over time. And, after designing a research approach that seems to incorporate the program's conceptual framework, researchers often must "tweak" it until it fits the circumstances of the study at hand.

One researcher who used a theories-of-change approach to study 20 family support programs found that it lent precision to program efforts, as practitioners realized that some of their practices did not match their goals at all. The approach also gave the researchers and practitioners a chance to grapple with tough program design and measurement issues, which established a more collaborative relationship.

Despite the benefits, focusing solely on theories of change presents several challenges to researchers and evaluators:

- A theory-based approach can complicate evaluation design because programs must accommodate the goals and objectives of diverse stakeholders (although the process of negotiating these compromises can add depth to research).

- In a context in which there may be separate theories of change at different levels of governance, it often is not clear whether a single theory exists or who "owns" a theory of how multiple initiatives fit together. Evaluation of the success or failure of a theory of change depends on whose theory is used.

- Simply understanding the theories of change that exist does not indicate how to conduct an evaluation.

Participants suggested that evaluators pay attention to the program funder's theory of change, in addition to that of the site. In particular, evaluators should determine the extent to which the funder's selection of sites is grounded in a theory of change because site selection can have a major impact on a program's success. Finally, participants agreed that evaluators should consider the conceptual frameworks that they use to examine programs, as well as the relationship between these frameworks and the evaluation.

Clarifying Attribution of Outcomes

Cause and effect can be especially hard to measure in research on comprehensive initiatives. The complexity of these initiatives and of the contexts in which they occur often makes it difficult for evaluators and researchers to establish causal relationships between program inputs and participant outcomes. Client conditions and the conditions of families and communities are closely interrelated; data on sites or participants, if analyzed in isolation, cannot prove that a particular outcome is the result of a particular intervention. A control or comparison group, which could show causality, may not be available for every comprehensive initiative. And data collected through qualitative and quantitative methods may indicate different (even contradictory) causal relationships.

Quasi-experimental techniques may help evaluators control contextual factors in order to establish causal relationships, suggested discussion leader Lynn Usher. For example, in Missouri several change efforts occur simultaneously but have diverse impacts across the state. A Family Investment Trust, funded by the Casey Foundation, operates as a state-level change agent. Community partnerships sponsored by the Kauffman Foundation act as local governing entities for selected communities and are designed to serve as vehicles for neighborhood change. Caring Communities, an effort funded by the Danforth Foundation, delivers comprehensive services to communities through schools. State leaders also are in the process of developing a new child protective services system, developing a mental health managed-care system, and overhauling the state's definition of child abuse and neglect.

In such an environment, Usher advised combining data collection on programs, synthesis of information across programs, and attention to system-level change. An evaluator might compare observations of outcomes in two communities over time to assess similarities and differences and then use the information to figure out what was done differently in each community that might have had an impact. Look for "counterbalancing indicators" to give a true sense of what is happening, he suggested; for example, in an evaluation of child placement services compare family reunification and program re-entry rates.

According to Usher, such a research approach results in common sense information that is easy to convey to communities. But he also warned against ascribing characteristics of a group or community to individuals within the group. Even when a systemic outcome appears to be positive—for example, a reduction in the out-of-home placements of children—individuals may suffer, such as when a child is kept in a dangerous or abusive home.

Understanding the Reality and Impact of Context

Because community context has a major impact on outcomes, researchers and evaluators must develop better methods for gathering background information on communities. "We have to put [context] in our summaries and our conclusions, and acknowledge [its] impact. Leaving it out isn't honest," said one participant. Context is an important factor because of (1) the impact it has on the actual programs being evaluated and the recipients of program services, (2) the impact it has on the speed and degree to which changes can occur, and (3) its ability to create organizational barriers to research. For example:

- Multiple reform efforts in Missouri generate interplay among policies and projects. In this environment, incorporating the context of reform into an evaluation enables researchers to tap into what the community identifies as its own strengths and needs and to determine which interventions are most responsive to local context.

- The school environment, particularly in traditional communities that have not attempted school restructuring, can either assist or challenge comprehensive school-based or school-linked initiatives. Some schools, like other established institutions, view the new kinds of relationships inherent to comprehensive service delivery as a threat to the school's established order. Other schools are

willing participants in change. A project's impact on the extent or nature of change in school culture may depend on contextual elements, such as whether the school sees itself as a host to outside services.

- Researchers have found that children in low-income neighborhoods are less likely to experience abuse and neglect when strong social institutions are present in their communities—but it can be difficult to define what constitutes a strong institution and whether that institution has had an impact.

The existence of multiple initiatives in a given community makes it difficult to determine which effort is responsible for results. Although the varied role of community factors in comprehensive initiatives makes it hard for evaluators to generalize findings across communities, incorporating community context into an understanding of local interventions helps researchers recognize the idiosyncratic nature of comprehensive services for children and families, one discussion leader said.

As Usher noted, the interplay among contextual factors makes evaluation of comprehensive community initiatives more a process than an event. In particular, said another presenter, evaluators and researchers should take into account the impact on a community and program of (1) multiple initiatives, (2) the expectations of key policy makers, and (3) the expectations of funders. For example, some foundations unrealistically want huge changes for a minimal investment; others provide "glue money" for local partnerships to leverage change. The apparent failure of a program may not be the community's fault—it may be the result of unrealistic expectations by the funder.

Ethnography as a Tool for Gathering Information on Context

Ethnography—research that combines unstructured interviews and on-site observations conducted intensively over a period of months—offers an important tool for gathering contextual information. The rich information provided by ethnographic research often corrects misperceptions and allows researchers to see situations from new perspectives; it also offers a "tangible, practical, and humane" research approach, explained discussion leader Susan Greenbaum.

The process of conducting ethnography helps build collaborative relationships between researchers and practitioners and supports the concept of the researcher as learner rather than outside expert, Greenbaum said. Ethnographic data collected early in a study can help researchers focus their research questions and determine what to include in quantitative measures. As one participant who had used ethnography explained:

> Through close, minute observations we began to understand and unpack trajectories of the initiatives. We began to see what worked and what didn't. With that information, we improved the quality of [our surveys].

Although traditional ethnography requires significant investment in time, staff, and funding, researchers can adapt certain ethnographic techniques and use them in isolation or in combination to enhance other approaches. For example, one participant suggested that it would be more cost-effective to conduct ethnographies before sites are selected to provide background for subsequent research—although this strategy would increase the cost of site selection.

If researchers do not have enough time to immerse themselves in the field, as a traditional ethnographer would, community residents can be trained to conduct ethnographic studies. "They already know the questions and have established a confidence within the community to get the data. We need to build that training into our design and capitalize on those resources," suggested one participant.

Researchers can also tailor ethnographic studies to address a single aspect of community life, rather than a broader cultural experience, or adopt specific ethnographic tools such as key informant interviews or physical surveys and observations.

Understanding, Acknowledging, and Incorporating Program Participants in Research and Evaluation

Better relationships between program participants and research and evaluation will require (1) more collaborative roles for researchers and evaluators, (2) cultural competence in the research and evaluation process, (3) the early involvement of evaluators in planning programs,

(4) information sharing and other efforts to build local capacity for research and evaluation, and (5) an understanding of the impact that evaluation has on program participants and communities.

Researchers and Evaluators Must Develop More Collaborative Roles

Conference participants unanimously called for evaluators and researchers to develop more supportive and interactive roles in their relationships with program sites, practitioners, and consumers. In particular:

Researchers and evaluators should be partners and collaborators with practitioners and other community stakeholders. Although evaluators have traditionally distanced themselves from practitioners, said Kagan, practitioners are actually ahead of evaluators in realizing that traditional service systems—and evaluations—no longer work. Practitioners also realize the need to reframe fundamental issues and have already begun the tough task of experimenting with new approaches.

Interactive relationships between researchers and practitioners are mutually beneficial. Through collaboration, practitioners gain a broader view of their program while researchers learn to clarify their language and make it more accessible to a larger audience, noted panelist Don Crary. Involving practitioners in designing and developing evaluation instruments encourages discussion between evaluators and practitioners regarding the substance of the evaluation. Participant involvement also can help researchers gain access to the community being studied. Foundations, which operate in both the practitioner and research worlds, can play a significant role in bridging the gap between researchers and practitioners and helping to "translate" their languages.

Ensuring that both insiders and outsiders—practitioners and evaluators—are involved in the evaluation process helps evaluators collect multiple points of view, makes data more relevant and useful to a range of interests, and increases investment on the part of the people who are affected by evaluation. Involving an array of sources also helps evaluators understand the context in which the evaluation occurs, observed panelist Weiss. The evaluator's role in collaborating with practitioners is to help various stakeholders negotiate their differences and

develop data that can lead to solutions. As researcher Tom Dewar said, reflecting on a recent project, "Our goal was to lessen the divide between camps...to jostle with the boundaries, not to resolve their questions."

Despite these benefits, participant involvement in evaluation takes time, makes research design more complex, and can sidetrack evaluations with programmatic issues. "I am asking us to get real in understanding how much engagement is appropriate, under what conditions, and to what end," cautioned one participant.

Collaboration With Practitioners Improves Evaluators' Access to Information

When Public/Private Ventures (P/PV) evaluated the Casey-funded "Plain Talk" teen pregnancy prevention program, researchers involved project staff and community members in developing and conducting a household survey of teenagers' sexual behaviors and knowledge. The site-based project staff believed that their community would benefit from the research training and from income generated from the survey project (P/PV paid survey interviewers). Casey and P/PV staff recognized an opportunity to build evaluation capacity within the community.

The local practitioners and residents helped P/PV create the survey, adding questions of interest to the community and identifying questions that might have cultural ramifications or raise privacy concerns. "The instrument development process was important because it began a dialogue around the substance of the survey," recalled P/PV evaluator Mary Achatz. "We addressed their concerns and adapted several questions....We benefited from their insights. We learned a lot from [this process]."

The on-site participants also helped P/PV identify community members who could serve as interviewers—people who could gain access to households quickly, make respondents feel comfortable discussing sensitive issues, and maintain confidentiality. P/PV and project staff jointly provided the interviewers with intensive training on survey and interview methods. The result: the community and project staff responded positively to the evaluation process, the information collection process was of high quality, and P/PV obtained rigorous scientific data.

Researchers and evaluators should serve as "containers and focusers of information."
U.S. policy makers have historically viewed social problems as "fixable," given the right solution; in this context, research was viewed as a means of influencing policy. Although there

was ample funding for such research in the past, policy-focused research became detached from program development, explained panelist Gary Walker. Today, policy makers increasingly focus on continuous improvement within multiple approaches rather than finding a single solution for fixing problems—and there is much less funding and political support for research.

In this context, said Dewar and others, researchers and evaluators have a special role in "listening, convening, describing, reporting, and discussing." Evaluators should be credible witnesses and analysts of what progress toward a goal means and what goals and thoughts are driving stakeholders in the field. Through this role, evaluators create a climate of learning and encourage sources in the field to share their best insights and practices with researchers.

Researchers and evaluators should document strengths and the quality of life that exist, rather than just problems and deficits. A focus on crises and deficiencies tends to compartmentalize people according to their needs and does not contribute to long-term self-sufficiency, while attention to strengths supports more preventive approaches and solutions that recognize the personal and community assets that do exist. "We have to get beyond the Grim Reaper model of evaluators," urged Weiss.

Researchers and evaluators should be facilitators of learning by all stakeholders—not just by evaluators. The structure of evaluation must include ways for stakeholders to learn from each other, and evaluators should address the questions that this raises, said Weiss. "What does continuous improvement mean—how do people use information to improve practice?" she asked her colleagues. "Are we as evaluators prepared to help them do that? What does it mean for us as evaluators to do that? What does it mean to truly create a collaboration?"

Individual researchers and evaluators should be part of a larger, unified evaluation community. The new, experimental evaluations required to measure complex community-based initiatives will be expensive and difficult to conduct—and not every researcher may need to concentrate on this type of work. As one participant suggested, efforts to create a more cohesive "community of researchers" could promote collaboration among researchers and help allocate research responsibilities to evaluators with an array of talents, resources, and capabilities.

The Role of Cultural Competence in the Research and Evaluation Process

Many communities and programs harbor a "basic distrust of and ambivalence toward evaluators and the world of research," Cipollone noted, practitioners and consumers in many minority communities where comprehensive programs are located often "don't believe that evaluators of different racial and ethnic backgrounds have the necessary experiences and sensitivity to understand and effectively analyze the context in which they need to work." For this reason, evaluations should not rely solely on evaluators who represent the mainstream culture—or on information gathered solely from mainstream sources. Research conducted by culturally diverse researchers, incorporating the perspectives of the people being studied and asking culturally appropriate questions in culturally appropriate ways, can provide richer information, rectify misperceptions, and address the issues and the goals that communities care about.

Cultural competence is not merely a racial or ethnic concern, it addresses the intrinsic differences in perspective between any "outsider" (researcher or evaluator) and "insider" (subject of the research). For example, when a discussion leader described a research project designed to learn whether community family centers would produce better outcomes for children and families in a particular community, a participant suggested that the outside evaluator's assumption that a new service structure was necessary might be inappropriate; people inside the community might feel that adequate models for providing services already exist but are simply underfunded.

Without research and evaluation that seek and respond to the needs and circumstances of the program being studied, we risk perpetuating service systems that do not respond to the needs of communities—systems that, as Cipollone noted, are "too inaccessible, too expensive, too irrelevant, too fragmented, and too often delivered far too late in the game to do anyone much good."

Involvement of Evaluators in Program Planning Facilitates Research and Evaluation

The evaluator should be involved in a program's conceptualization and throughout the site selection process. This provides an opportunity to raise questions and issues on a program's

BEST COPY AVAILABLE

direction that can have an impact on its evaluation. For example, as one presenter from Kentucky explained, the fact that the Kentucky State Legislature did not mandate or consider a formal evaluation as part of that state's sweeping 1990 education reform act has complicated future efforts to develop a comparative look at the reform's results.

Building Local Capacity for Research Has Many Benefits

Building on-site capacity for research makes evaluations more relevant, increases respect and reduces hostility between practitioners and evaluators, and makes the information collected by evaluation more useful because it can be used for planning, practice, and to improve communication with members of the community. Information sharing is an important part of capacity building because it helps build a foundation for data analysis in communities and encourages local participation in evaluation. In the Plain Talk evaluation by Public/Private Ventures described earlier, for example, the community insisted on receiving the raw data and learning how to analyze survey results. After participating in the evaluation, community members felt that they owned the data, said evaluator Achatz. By providing the sites with training in data analysis, evaluators gave practitioners and project leaders the necessary tools to adapt the evaluation product for local purposes.

Training local practitioners in evaluation measures also builds capacity by making it more likely that they will use self-assessment techniques, suggested panelist Sandy Weinbaum. However, the technical assistance provided by researchers and evaluators to build local capacity requires a time investment.

Participatory Assessment Gives On-Site Practitioners a Chance to Develop Evaluation Skills

- A three-year assessment of youth-serving organizations in New York City, conducted jointly by evaluators from AED and frontline project staff, helped after-school youth projects use findings to influence planning and practice and to communicate with stakeholders in the community. First, AED invited all organizations that received city funding for youth services to identify the issues they thought were important in reaching community members and capturing the successes of their projects. The answers provided AED with a sample of the kind of questions that the study should address. Next, AED requested proposals, eliciting 15 applications. AED accepted six proposals from this pool.

  AED developed 30 hours of workshop training for project staff. Workshop topics focused on demystifying the process of evaluation, helping staff identify goals and objectives as well as practices that support them (i.e., eliciting their theories of change), and designing and conducting an evaluation of one aspect of the participants' projects.

  "We looked at the different audiences [for evaluation data]: what information do you need to collect, how do you pose questions...and then taught them how to actually use these strategies," Weinbaum said. Strategies included surveys, observation and interviews, and focus groups. At first, project staff did not want to let the other workshop participants know their program's faults. Once they realized they were all trying to help children and families, however, participants became more collaborative.

  AED and the participating organizations held focus groups and helped participants establish information tracking forms. The organizations then taught their staffs how to use the forms. According to Weinbaum, the participatory process raised questions never before asked within some organizations and gave participants a lasting means of improving their learning.

Research and Evaluation Has Practical and Political Impacts on Practitioners, Consumers, and Communities

Researchers and evaluators have a responsibility to realize that the subjects of a study have a personal interest in research and are affected by evaluation in several ways. As panelist Dorothy Coleman explained, program staff and those they serve care what their community thinks about them; they have an investment in their local reputations, and the results of research can change these reputations. Second, although program leaders and staff are eager

for assessment information, they do not want the data to expose or exploit individual children. Third, program staff want to know what impact the evaluation outcomes will have on the community, families, individuals, and the program itself. Fourth, program leaders and staff want to know what investment or involvement evaluators—and the evaluation itself—have in the community.

Participants agreed that evaluators must address these issues. For example, evaluators might measure their own attitudes, or review historical accounts by key players in the organization, to discern the changes that occurred in an organization as a result of the evaluation itself. Coleman—a practitioner and participant in a current evaluation—described being asked by a 17-year-old participant in her program to explain what the evaluation could tell Coleman that she didn't already know. "I said, 'What I've learned in the last seven years that I've been here is that working with families is like a game of poker,'" Coleman recounted, borrowing from a popular song:

> You've got to know when to hold them, when to fold them, when to walk away, and when to run. I know in working with this young lady, I saw something that told me to hold and to walk cautiously. I'm hoping the outcomes of [our] evaluation will teach us how to hold on a little longer and how to walk a little slower.

Researchers also must realize that reform initiatives are political processes that do not operate in a pure research environment. "Every time a researcher or evaluator comes into our city, talks to people, and issues reports...there are political consequences. And what is said is interpreted in political ways," observed panelist Crary. As a result, research has the potential to criticize or validate public expenditures, the performance of program staff, and the judgment of public officials. "Listen to your sites about what political impact you are having," he urged:

> When a site whose commitment is to a successful initiative feels there is negative fallout from evaluation, research loses access to good information and the work will be discredited at the local level.

Researchers must learn not to be ambivalent about the political process, some participants warned. Researchers and evaluators want research to be rational, sequential, clean, and logical—but the issues that end up in the political process are complicated. As a result, noted panelist Ralph Smith, researchers may approach the political process as if there is something wrong with it, which is not necessarily the case.

Using Multiple Strategies to Improve Research and Evaluation

The use of multiple data collection and analysis strategies makes research and evaluation more powerful, comprehensive, and compelling, participants said. Different audiences for information require different types of information—and, consequently, different data collection strategies, ranging from interviews and observation to surveys and focus groups.

Multiple research techniques are increasingly important as researchers begin to incorporate complex factors such as community context into evaluations. For example, data collection methods such as ethnography, participant interviews, and the use of management information system (MIS) data complement each other by providing different types of information on the many needs and strengths of a community. While MIS data show the number of people receiving a service, ethnographic interviews of service providers can reveal the way in which service use has changed over time, and interviews with service recipients can reveal the contextual factors that create demand for the services—or flaws in the services themselves.

Mixing strategies is also useful for gathering information attractive to policy makers, who like interesting, rich stories but also want hard numbers. Using multiple techniques enables researchers to collect compelling stories but guards against criticism that ethnographic or anecdotal data is not "hard" enough. "I agree with the power of stories...but I also think that stories are dismissable. If you have stories with data, then you have the best of both," Greenbaum explained.

Whatever the mix of strategies, research and evaluation should take risks by offering new ideas and provocative analysis, observed Cipollone and others. Evaluators may not

necessarily throw out all of the traditional approaches but must at least explore and experiment with new techniques, agreed panelist Weiss.

Researchers and their funders must also realize that the kinds of shifts in focus advocated by those on the cutting edge of evaluation—ethnographic methods, longitudinal studies, local capacity building, and other approaches—require a greater investment of time and resources than many evaluations currently allow. As Kagan noted, evaluation funders—including the federal departments of education and health and human services—must be willing to take risks on innovative evaluation as well as on innovative programs. "This is not just a foundation responsibility. We're not going to get over the hump without investment [from all sources]," she said.

Using Research and Evaluation Information to Improve Programs and Policies

In order to support continuous improvement and build a strong case for increased funding for comprehensive services, stakeholders must learn better ways to use the data and findings collected by evaluation, participants agreed. Researchers and evaluators must address how people can use information to improve practice, whether evaluators are prepared to help practitioners use the data, and what implications the evaluator's involvement in information use has for his or her role as an evaluator. This means (1) understanding how and when research influences policies and (2) producing and promoting useful information.

Understanding How and When Research Influences Policies

Research rarely affects public policy directly; policies are influenced by a combination of information, ideology, and the personal interests of policy makers, observed discussion leader Fran Jacobs. The challenge is to determine when facts matter and when they don't—and, when they do matter, to find ways to encourage deliberative decision making, said panelist Smith.

Most political fights are concerned with values, not research—and research does not help fights about values, said panelist David Ellwood. But research that begins with agreed-upon values can focus on the best strategies for achieving goals. Researchers who begin with

"what do we all believe, and what are we trying to achieve" and are explicit about values will gain credibility and the attention of policy makers, he said.

There are times when policies are so entrenched in values that research will not make a difference. But research will make a difference when very strong results exist that have a bearing on the situation at hand—for example, unambiguous and compelling information on methods that would reduce teen pregnancy. Research can also be used to shape politically driven policies (e.g., to determine when to stop the clock on time limits under the Worker Responsibility Act). For these reasons, research should be based on real and relevant issues, concepts, and goals, Ellwood said.

Producing and Promoting Useful Information

The current era of scarce resources for programs, services, research, and evaluation places a priority on the production of information that is both usable and used, participants agreed. Effective types of evaluation information for policy makers include baseline data, longitudinal studies tracking long-term effects, comprehensive evaluations of multi-faceted initiatives, and data that appeal to policy makers' economic motivations as well as to their consciences.

Research information targeted to legislators should be simple and direct, preferably containing simple concepts to which constituents will respond. Policy makers "don't like to be yelled at, don't want to be seen as corrupt, and want simple information," summarized Jacobs.

Research should elucidate how real people live and what their issues are; it should help policy makers understand how they can make a difference in real concerns. For example, a participant who evaluates family support programs traces family histories through major events and examines the responses of systems and services to family needs. "You see pathways and responses, and how the system can fail to respond," the researcher explained. "We track the most troublesome children, like those who cost the city $1.5 million alone, and we look at their pathways. [When] you tell those stories to legislators, you personalize it and build it case by case."

Evaluators Can Use Cost Data to Show the Economic Benefits of Comprehensive Services

An evaluation of a range of community outcome indicators in one county, conducted by Bruner, included an economic impact study that showed the potential economic benefits of developing community-based family centers.

Bruner first examined outcome indicators such as the percentage of low-birth weight babies and the percentage of families living in poverty. He then assessed the cost of services to these community members across multiple human service systems—welfare, Medicaid, food stamps, child and youth services programs, judicial services, etc. When Bruner compared the amount invested in these services in this high-need community to that invested in average communities, he found an estimated economic difference of $563 million.

Bruner's cost assessment helped to make a case for devolving human service delivery to neighborhood community centers as part of an economic development strategy. "We saw family centers as just a piece of the answer," Bruner said. "We tried to say, 'How much are we spending now within our system to do some of this work? How much could contribute to [a more] holistic approach?'"

Getting Policy Makers to Use Research and Evaluation Data

Researchers and evaluators should be proactive in getting policy makers to use the information they produce, participants said. In particular:

Researchers must simplify and distill what they know and present it in much more compelling ways because the way in which evaluators frame the public discussion of comprehensive services has an impact on public support. Instead of hedging their bets about what they *don't* know, researchers must be clear about what they do know and say it in a very straightforward way, said Cipollone, Crary, and others. Findings should be presented in plain language, common to both practitioners and researchers, that will promote learning rather than simply show off expertise; the content should reflect local context.

Although researchers may feel as if they are "lying" if they simplify information, there is usually a powerful and basic theme, fact, or message that can be used to focus the information, noted Ellwood. "You can make clear your message without damaging the

complexity of the problem or issue, and it makes your argument more powerful," he advised. Further, research should be presented as definitive, even if related issues remain unresolved.

Recognizing the power of the sound bite in public policy making, some participants advised researchers and evaluators to use language that respects "the public mood." For example, the term "home building" is more publicly acceptable than "family support." Similarly, researchers should reconsider "the preventive message" because "people do not want to pay for what should not happen anyway," suggested a participant.

Researchers should use real language that reflects local experiences, not analytic jargon or abstract discussions of research categories. As Dewar recalled, for example, the strong programs in a study he conducted did not think of themselves as "models" or "partnerships"; they viewed themselves as participating in a process of "negotiating working relationships." "We tried to use language and content that reflected local history and context," Dewar said.

Researchers must do a better job of getting what they know into local public discourse so that it shapes public discussion and sends a message to policy makers directly from the public. Panelists advised researchers to identify advocates who can speak (in sound-bite form) about programs and necessary changes without trivializing the issues. Researchers also should identify policy makers' key legislative staff and communicate regularly with them because they are critical to maintaining and changing programs, a panelist advised.

Researchers must be willing to jump into the fray of advocating particular solutions, rather than letting public policy drift in other directions and then trying to bring it back. Unless researchers apply their findings to possible solutions, policy makers may respond without appropriate input from the field, warned Crary:

> Public officials will respond to a crisis—whatever the crisis is perceived to be by the media and the public. Without us entering that dialogue, they will look for and find quick and obvious solutions....Policies will be set and budgets allocated in those directions. If we believe what we're about...we have to be clear about how [to] address problems and argue for solutions.

Researchers should not be afraid of "shuttle diplomacy," advised panelist Penny Sanders. If you have a chance to talk with a key policy maker, grab it—but make your conversation to the point.

Researchers must realize that providing information to policy makers presents a good news-bad news dilemma: research and evaluation may identify flaws, but politicians don't want anything but good news. This political reality means that researchers may face a tough sell in proposing to continue or correct a program that does not show unqualified success.

Using Information to Give Policy Makers a Realistic Sense of the Impacts of Policies and the Viable Pathways to Achieving Better Results

From Cipollone's opening remarks to Schorr's closing address, conference panelists, discussion leaders, and participants reiterated the need to use information more effectively to educate and motivate policy makers. "Change can only be effective and durable if it can successfully bridge the real worlds of policy and practice in a meaningful way," advised Cipollone. Relevant information on impacts, outcomes, and options can have a significant effect on policies, he continued:

> We believe that we can convince governors and state legislators concerned with issues related to urban children's mental health that it is possible to integrate state agencies, move decision-making to communities, and place more authority for resources in the hands of neighborhood residents if they can be shown—through the real-life examples of similar states—that such efforts improve service effectiveness, enhance service efficiency, and enforce greater accountability for achieving improved outcomes.

By drawing the connection between policies and impacts, researchers can show the ineffectiveness or even destructiveness of current systems, observed Ellwood. For example, a study focused on changing the culture of welfare offices showed that welfare recipients expected the offices to be responsive sources of assistance—while in reality, the bureaucratic nature of the welfare system reduced the offices to a check-writing and eligibility-verification role.

Similarly, providing information on outcomes and results has the benefit of "exposing the sham in which human service providers, educators, and community organizations are consistently asked to accomplish massive tasks with inadequate resources and inadequate tools," Schorr noted. "Attention to results forces the question of whether outcome expectations must be scaled down, or interventions and investments scaled up to achieve their intended purpose," she said.

But providing the "hard" information on impacts that lawmakers seek while acknowledging the unique demands of complex, community-based or school-linked initiatives is a challenge for researchers. An evaluator of a multi-site, school-linked service program described her difficulties in balancing the broad outcomes articulated by the program (for students to be healthy, happy, productive citizens who complete their education) with data on teen pregnancy, school dropout rates, and other issues. Evaluators are unsure of how to talk about the program's global, general outcomes with legislators or funders who seek more definitive results. One option has been to compare the costs of the program to the social costs of housing youth in foster care, detention facilities, prison, or other out-of-home options—an approach similar to the economic impact study described by Bruner.

The Challenge of Using Evaluation Results to Influence Lawmakers' Policy Decisions

As evaluations focus on the broad goals and theories of change expressed by programs, researchers struggle to satisfy the needs of legislators or funders who seek "harder" results. "I don't think policy makers recognize who kids are and how intense and complicated their needs are," complained one participant. "Legislators say, 'If you don't have data and proof, don't even talk to me about any program,'" agreed another.

Despite the frustration of meeting lawmakers' needs, researchers and evaluators are beginning to understand the roles they must play in making information useful and usable for policy makers. "[T]he moral underpinnings for social action...are not powerful enough today,

in the mean and cynical closing years of the twentieth century, to sustain what needs to be done on the scale at which it needs to be done," warned Schorr:

> In this time of pervasive doubt, we have to be able to provide hard evidence that investments are achieving their purpose and contributing to long-term goals that are widely shared, if we are to have any hope of obtaining the magnitude of public investment that is required.

Conclusion

Although conference participants recognized that there are no clear, easy solutions to the challenges of making evaluation useful for programs and policy, they generally agreed on the issues and needs involved in improving the collection and use of information about comprehensive community initiatives. Participants also agreed that although the process of improving research, evaluation, and data use presents challenging issues, it is worth the struggle. Many public policy makers are seeking good information to inform good decisions, and researchers can find ways to provide this input. "If we can be clear about information and show that it provides answers, we can convince them," said one panelist.

Finally, participants responded to Kagan's call for researchers to "get smart about what doesn't work, get real about new approaches, and get going." Although the challenge of creating new evaluation techniques and using research information more effectively requires a degree of humility, Kagan said, "humility should not stand in the way of action."

Appendix A:

Dinner/Welcome Speech
Tony Cipollone, Associate Director, Education Reform, Research and Evaluation
The Annie E. Casey Foundation
Wednesday, September 27, 1995

When we put together the guest list for our first evaluation conference in 1994, I remember being struck by the impressiveness of the list. I remember saying to someone, "Do you really think we'll get all these people to come?" I was not only pleased by the attendance at last year's conference, but overwhelmed by the level of participation and nature of the feedback that we received. We were so impressed, that to Donna Schmidt's chagrin, we decided to do this again. Of course there were some who felt that we had to do another conference because in my concluding remarks, I said that things went so well that we'd make this an annual event.

But reasons aside, let me take this opportunity, on behalf of the staff of the Annie E. Casey Foundation, to welcome you to Baltimore and to our second annual evaluation conference. I know I also speak for Doug Nelson, our Executive Director, who was unable to join us this evening, when I say that we are once again buoyed by the positive response to this meeting. We are honored that so many thoughtful people, engaged in such good work, have chosen to spend a few days with us as we grapple with what we think are a set of provocative issues.

Many of you were here last year, when our theme was "Reforming Systems, Reforming Evaluation." It was, in many respects, an interesting opportunity to test out some new ideas and directions concerning the ways in which the field of evaluation might begin to change and evolve so that the frameworks we use, the methodologies we employ and the analysis we bring to this enterprise could be more responsive to the complexity of comprehensive reform efforts.

I think that it's safe to say that we still have lots of issues, directions and debates to resolve on that front and, in fact, probably could have crafted this year's conference as a sort of second round version of last year's meeting. But there have been some important events between then and now that I think helped us move toward a different agenda. Specifically, the most outstanding events have been the federal policy changes that have taken place since last September. Winds have shifted, policies have changed, priorities have been replaced. A new national agenda has evolved. And I think one factor that struck many of us over the course of the last several months has been the degree to which the prevailing political rhetoric seems to have been fueled more by perception than fact—by anecdote rather than evidence. In large measure, the field of research and evaluation has, by our observation, weighed in very little in the development of the current New Federalism.

This observation led us to conclude that it might be worthwhile and helpful to examine the issue of utilization and collectively discuss issues that can more readily—at least in the context of our work—make a contribution to the development of effective policies and practices, a contribution to the fostering of stronger reforms that can better address the needs of kids and families.

On a lighter note that goes beyond the wake up call we all got around the new national policies, it might also be appropriate to defend the theme for this year's conference on the basis of what we at the Annie E. Casey Foundation describe as the "Middendorf Challenge."

The Middendorf Challenge is named after one of our most prominent trustees--Frank Middendorf—a now retired Director of Operations for the United Parcel Service. Frank, who I admire greatly for his ability to cut to the chase on just about every issue from baseball to Bosnia, is, for many of us, the consummate steward of Annie E. Casey Foundation resources. Frank Middendorf is a practical man who deals in practical issues, so that when any of us, from Doug on down, weaves any sort of intricate framework, theory, or elaborate funding idea that we think may be brilliant or show off our diligent work, we know—we absolutely know—that Frank will be there, holding our feet to the fire, applying the Middendorf Challenge in response to our presentation. Put simply, the Middendorf Challenge translates into the following statement and question: Well, all that sounds real good, but what's it going to do for kids?

And when we apply the Middendorf Challenge to the world of research and evaluation, a similar but slightly different question gets asked, which goes something like this: Gee, all this information is great, but do you really think that anybody is ever going to use this stuff? And, Middendorf or no Middendorf, that seems, at least to us, to be exactly the question we all need to be asking and is probably as relevant a backdrop as any for this conference.

Given this, I thought it useful, and, I hope, relevant, to spend just a few minutes this evening helping to frame the utilization issue a bit—what it means for us at the Annie E. Casey Foundation, why we think it's important, and why achieving it may be one of the biggest challenges we face.

To do justice to this issue, it may be important to begin by laying out some of the basic reasons behind our decision to even fund evaluation in the first place. Why, given the multitude of issues, programs, and needs that we may have an opportunity to influence, have we chosen to spend several million dollars over several years on an array of note taking, tape toting smart folks who don't always provide us with good news? Why, some of our critics might ask, do we devote so much time, energy, and resources to an area that in many foundations and organizations devoted to children, gets relatively short shrift?

Clearly, one important reason is our belief that evaluation serves as a strong accountability tool. Evaluation, when done well, can help the Foundation better answer and understand the degree to which our investments represent good judgments about the

organizations, communities, and people in which we place our confidence and dollars. Evaluation, when done well, can also, importantly, tell us at the Foundation what we need to know about the soundness of our theories, the practicality of our initiative policies, and the degree to which we have crafted the right kind of working relationships with grantees.

For example:

When done well, research and evaluation ought to tell us whether we can really affect teen pregnancy rates through strong community-generated messages about sexual behavior, the cultivation of stronger relationships between and among youth and adults, and more relevant and accessible community services for teens.

When done well, research and evaluation ought to inform us about the viability and potential of using strong, experienced, community-based organizations as catalysts for the real transformation of neighborhoods.

When done well, research and evaluation ought to tell us what community-based alternatives to secure juvenile detention look like, how they're created, and whether they serve to reduce the population of overcrowded juvenile detention facilities without sacrificing public safety.

When done well, research and evaluation ought to tell us about the appropriateness of our planning timelines, the relevance of our technical assistance, and the degree to which we have built true partnerships with our grantees.

But beyond accountability, it strikes us that there's another compelling reason for a serious and sustained investment in research and evaluation activities. We invest in research and evaluation because these activities are, in a very real and very practical way, an important component of the Annie E. Casey Foundation's theory of change.

As a Foundation, we are unabashedly and unapologetically committed to changing and improving the life outcomes for our nation's most disadvantaged children and families. It is our mission, our motivation and our message to the world. We stand for kids and families. And driving our investments, and undergirding our actions, is, we believe, a strongly reasoned orientation about the problems facing America's poorest families and a theory of change that we have sought to test in the variety of initiatives we've undertaken over the last five years or so.

It is an orientation and theory firmly rooted in the belief that outcomes for kids will not change unless and until our country, our states, our counties, our cities and our communities foster a set of fundamental, comprehensive, and durable changes in the multitude of systems that currently operate in our poorest communities. It is an orientation, rooted in belief and borne of experience, that our nation's poorest families live in communities devoid of opportunity and littered with service systems—systems of health, education, emotional

support, juvenile justice, and jobs, among others—that don't work effectively. Systems of service that can best be characterized as too inaccessible, too expensive, too irrelevant, too fragmented, and too often delivered far too late in the game to do anyone much good.

The combination of these factors contributes greatly, in our estimation, to a cycle of ineffectiveness in combating what is now an intergenerational cycle of poverty characterized by families and young people with too little hope, too little opportunity, and far too much cynicism, pain, and despair.

But we stand fast on the belief, the theory, the assumption, that the conditions that characterize our service systems and our communities are not irreversible. We are driven by the conviction that change is indeed possible, that communities can prosper, that families can thrive, and that children can learn and develop when neighborhoods are supportive, sustaining, and served by systems that are relevant, respectful, and rooted in the communities they serve.

Our investments—whether in New Futures, Plain Talk, Family to Family, Mental Health, Juvenile Detention Alternatives, Rebuilding Communities, or Education—operate within a theory of change that goes something like this:

The Annie E. Casey Foundation can, through strategic investments in awareness building, capacity development, program demonstrations, and research and evaluation, help move currently dysfunctional systems in new and more productive directions. Directions that are characterized by greater collaboration that can foster the integration of services, decategorization that provides greater flexibility around services and resources, decentralization that invests greater authority for services and finances in neighborhoods and with those who are closest to kids and families, and meaningful incentives and sanctions that can indeed promote greater accountability for achieving better outcomes on behalf of kids and families.

Moving toward such new systems of service and support requires and demands much in the way of preconditions. It requires, as we've seen over the course of the last few years, tenacious leadership, moral persuasion, a strong sense of the technology of innovation, more than just a little political power, some new money, and some luck. And we are convinced that it requires and can benefit from accurate, relevant, and compelling information—the kind of information that we want and desperately require from our research and evaluation efforts.

For example:

We believe that teachers and parents in schools will make better decisions about instructional and organizational practices if they better understand the degree to which academic failure, suspension, and dropout rates disproportionally affect children of color and those who are poor.

We believe that communities who seek to involve more adult males in their efforts to combat teen pregnancy can benefit from strong case studies of communities who have successfully engaged males in their Plain Talk efforts.

We believe that states can and will make better decisions about the expenditure of resources and efforts if they better understand the nature of out-of-home placement rates within the child welfare and foster care systems.

We believe that we can convince governors and state legislators, concerned with issues related to urban children's mental health, that it is possible to integrate state agencies, move decision making to communities, and place more authority for resources in the hands of neighborhood residents if they can be shown—through the real life examples of similar states—that such efforts improve service effectiveness, enhance service efficiency, and enforce greater accountability for achieving improved outcomes.

In short, we believe that research and evaluation can be, should be, must be a critical, compelling, and integral component of comprehensive reform strategies. Good research and evaluation has the potential, if properly used, to increase the power of those living in, working in, and working for traditionally disenfranchised families and communities. Research and evaluation are a conduit to information, and information, as we all know well, is power.

But let's be real about all this. We know, and we know well, that the power of information—the power of research and evaluation—is like the power of a tool or like the power of an athlete, in that it will remain dormant unless it is perceived as being useful to a given situation.

But what makes for research and evaluation that will be used by those who might create, sanction, and sustain changes in policies and practice? What is the litmus test we need to meet? For starters let me offer three simple, common sense characteristics that come out of our own dealings with grantees.

First, if evaluation is to be used, it needs to be presented in an interesting and provocative manner. By this, I mean that it needs to be clear, it needs to be concise, it needs to be conveyed in a way that speaks plainly to people. It needs to relate real examples of change using the voices, words, and experiences of real people. And, when appropriate, it needs to take risks by offering new ideas and provocative analysis. As we used to say when we were smart-mouth kids: "Tell me something I **don't** know."

Second, if evaluation is going to be used to effect change, then it has to answer questions and address issues that people really need to know something about. That is—it needs to be relevant to the work that people in states, cities, and neighborhoods are actually trying to accomplish. For example, the relevance of an evaluation of local governance structures may be a function of its ability to not just report on whether new governance structures work, but to examine and discuss wider issues—the kind of leadership necessary to

sustain local governance, the way successful collaborative decisions are made, and the navigation of relationships between new governance structures and old structures with more history and formal authority.

Third, for evaluation to be truly useful, it's going to have to be able to transcend and speak to—in an equally compelling way—diverse audiences. A common thread that weaves through each of our comprehensive change efforts is an emphasis on fostering change across different environments—political chambers, state agencies, city government, community-based organizations, schools, and the community room of the local housing project. Specifically, we believe that change will only be effective and durable if it can successfully bridge the real worlds of policy and practice in a meaningful way. Similarly, for evaluation to be utilized, it must not only be relevant and speak in an interesting and provocative way, it must do so in a way that grabs the attention of the multiple audiences that operate in these venues: politicians, policy makers, practitioners, parents, and the public at large. Given this, I am particularly looking forward to Lee Schorr's remarks this Friday, because I think that she has hit this issue exactly right: How **do** we help the day-to-day work of Sister Mary Paul, Geoff Canada, and Otis Johnson, while building a case that is compelling enough to convince Pat Moynihan, Newt Gingrich, and the American Public?

Clearly, these are criteria that may be simple to describe, but—I know from our experiences at the Foundation and I suspect from your own—difficult to address. And, as we know from the thoughtful and articulate work of folks such as Anne Kubisch, Charlie Bruner, Joy Dryfoos, and the other individuals who have been kind and gracious enough to agree to lead some of our sessions over the next few days, the challenges inherent in crafting research and evaluations that meet these and other criteria and that are powerful enough to influence policy and practice are many and significant.

But in the spirit of the day (or evening), let me quickly leave you with four challenges that I'll pose in the form of questions. I raise them not just to be provocative, but because they get raised often as we interact with staff, evaluators, and grantees; and quite honestly, we could use some good thinking in these areas.

First: how do we craft useful evaluations when we often work in environments that may have a negative history with respect to research and evaluation? Simply put, there is, in many of the environments that we need to successfully reach, a basic distrust of and ambivalence toward evaluators and the world of research. People are often afraid of what you'll say, reluctant to put in the necessary time because they see no payoff at the end, and don't believe that evaluators of different racial and ethnic backgrounds have the necessary experiences and sensitivity to understand and effectively analyze the contexts in which they need to work. In short, how can evaluation be useful in environments where it may be, at best, merely tolerated?

Second: How can we help people develop the skills and experiences to effectively use even the best and most relevant evaluation information? I think we all know that the

relationship between comprehensive change efforts and research and evaluation is not the least bit akin to that between old baseball players and the baseball field in *Field of Dreams*. If you build it, they may come; however, if you research and evaluate it, even when it's relevant, interesting, and provocative, they may not use it—not because they don't want to, but because they may not know how. So as we think about how to create evaluations that are more useful, we also need to think hard about the kinds of skills and experiences that foster successful utilization. Clearly, one important issue in this area is a need to come to grips with the role that the evaluator plays in such a process.

Third: how do we push ourselves—and this is relevant to both the Foundation and evaluators—how do we best push ourselves so that we can effectively capture interim benchmarks of change that tell us all whether we are on track to achieving the long-term outcomes we seek. If there is validity in the connection between utilization and relevancy of evaluation information, then we all need to find better ways of conveying the short-term progress of long-term change efforts because it is that progress that offers the most help for the ongoing work of our grantees. I think that we may be making some progress in this area, but offer it as a major issue to be addressed.

Fourth, and finally: if we are to move in the direction of fostering greater utilization of research and evaluation, how do we think about dissemination strategies that can effectively reach and transcend the varied audiences I noted earlier? What are the right forums, programs, publications, and technologies we need to enlist to make a truly discernable difference in the way we use information around comprehensive reform?

While there are clearly no easy and definitive answers to these and the other challenges to utilization that will be raised over the next few days, my hope is that we can, given such impressive company, shed some light on how we think about them.

In closing, I must admit that I feel a much different sense of urgency than I felt last year at this time. As I noted earlier, we have, in the last 10 months, been buffeted by lots of rhetoric that implies, most positively, that we don't know what works on behalf of poor families, or most negatively, that nothing works. And I must confess that it bothers me tremendously that as a nation, we are not able to make timely, well-thought out, data-driven decisions on behalf of poor kids and families. My hope is that through efforts like this conference and the good work that is already going on in the various venues that are represented in this audience, we can, over time, make a meaningful contribution to the debates, the decisions, the practices, and the policies that make a real difference.

My hope is that we can, over time, face the Middendorf Challenge with confidence and tell the world: Here is why all this stuff is useful, and this is how you use it to make a difference for kids and families.

Please enjoy what we all hope will be a useful and interesting conference.

EVALUATING COMPREHENSIVE, COMMUNITY-BASED SERVICES: RETHINKING PURPOSE AND PRACTICE

Sharon L. Kagan, Ed.D.
Yale University

During the past several years, the evaluation of comprehensive, community-based services has become a central topic discussed at seminars, colloquia, and conferences, and addressed in numerous publications (Behrman, 1992; Bruner, 1994; Connell, Kubisch, Schorr, & Weiss, 1995; Coulton, 1992; Crowson & Boyd, 1993; Knapp, 1995). The purpose of this paper is to provide an action-oriented agenda for researchers and practitioners predicated on the work done to date. The paper suggests that while the state of evaluation methodology remains in need of refinement, we need to: (1) get smart (by acknowledging what doesn't work); (2) get clear (by examining the outstanding fundamental issues to be resolved); and (3) get going (by experimenting with new and promising approaches).

GETTING SMART: ACKNOWLEDGING WHAT DOESN'T WORK

For many years, practitioners have recognized that some fundamental practices were not working—top-down decision making, input rather than outcome-driven accountability, governance for the many by the few, uncoordinated versus linked services, a focus on the individual at the expense of the family and community. Energized by these concerns, the education and human service communities began to redefine their premises. Codified by some in the principles of the family support movement, by others in educational reform efforts, and by still others in social service and community regeneration efforts, new ideas about comprehensive, coordinated, community-based services took hold.

Adopting a similar stance, we in the evaluation community need to acknowledge that some of our conventions do not work, in part because of the inherent and complex nature of the new approaches being used, and in part because some conventional evaluation strategies were not working anyway. At least three lessons are worth noting.

Get Smart One: Acknowledge the mismatch between design and practice. Researchers and practitioners like the design tightness of conventional research strategies, yet we recognize that they are fundamentally misaligned with the flexibility inherent in comprehensive, community-based services. Control group design does not account for the "noise" in the field or for the inevitable contamination that occurs as the comprehensive, community-based service movement grows. New interventions proliferate offering alternative options to historically virgin communities. Indeed, the lives of participants are not static—as they were tacitly presumed to be. Random assignment, the *sine qua non* of experimental research, is foreign to the inclusive, on-demand quality of comprehensive programs. Units of analysis, once clear and distinct, have become murky and often difficult to discern. In short, what was difficult for practitioners to tolerate and researchers to accommodate is next to impossible, given the principles and trajectories attendant to comprehensive, community-based efforts.

Get Smart Two: Acknowledge illusive independent variables and shifting bottom lines. Knapp (1995) notes that in these new efforts, the independent variable often ceases to be a fixed treatment, giving way to a "menu of possibilities" (p. 7) that meet participants' changing needs. Difficult to hold constant for any single participant, the treatment also varies across participants. Gone are the days of the single treatment that remained constant over time. Indeed, the goal of many of today's comprehensive, community-based efforts is to force change. Mid-course corrections are necessary and desirable; they are not nuisances. Complicating the matter even more, many of the efforts are collaborative in nature, often expanding the treatment across agencies and services. What then constitutes outcomes and to whom are they attributable? Are the outcomes associated with changes in individuals? with changes in institutions? or with changes in the processes that transcend institutions (e.g., service integration, collaboration)? Who caused what for whom?

Get Smart Three: Acknowledge the importance of process. Because change is fundamental to the intervention and because so many parties are involved in these interventions, we need to focus more on the process and the process of change. How can different disciplines' views of the effort be incorporated into the design and evaluation? How can the perspectives of diverse participants be factored into the interpretation of results? How can especially sensitive interactions be captured, their impact understood and accounted for? Sometimes evaluations have failed to account for the nuances of these changes. On other occasions, the evaluations may be too short lived to capture durable change. In still other instances, changes in one element of the intervention may be counteracted by changes in other dimensions. However unwieldy, process studies are critical to replicability and to advancing our understanding of how to get things done. Sadly though, our evaluation psyche does not accord much credit to process evaluations or the systematic studies of change.

In sum, in getting smart, we need to acknowledge that many of our evaluation "sacred cows" were created to examine changes in individuals, rather than changes in systems, neighborhoods, or communities, the thrust of the current waves of reform. Gone, then, is the applicability of the conventional research paradigm. The real new learning is that we simply cannot fit the square peg of conventional evaluation into the round hole of comprehensive, community-based efforts.

GETTING CLEAR: REEXAMINING FUNDAMENTAL ISSUES

Having acknowledged extant conditions, evaluators—like practitioners—need to examine fundamental issues to move forward. In the field, the examination of issues took structural and attitudinal forms. The field asked itself what are the organizations, the governance apparatuses, and policies fortifying or obstructing the new intents: can this system be tweaked or dramatically reformed to accommodate community driven interventions? The field also asked itself what is the nature of prevailing attitudes toward change? How and can we move—given external realities of personnel training conventions, in-service capacities, and embedded values and beliefs—to a fundamentally different approach? The point is that

practitioners dug beneath the surface to identify and challenge their conventional thinking about deep-seated issues. And in so doing, their strategies became more clear.

Get Clear One: Outcomes. As suggested above, many comprehensive, community-based efforts have difficulty pinpointing the outcomes to be examined: are they child-based? family-based? community-based? Are they short-term, mid-term, or long-term? Are they dependent or contingent: should they measure behavioral changes in individuals or social indicators reflecting changes in communities? In getting clear on the dependent variables, an important exercise might be to discern what is not an outcome of the intervention. The temptation is to match the ambitiousness of the intervention with an equally ambitious evaluation. We need to be quite clear on the desirability of this, and perhaps elect to be more parsimonious in the definition and measurement of outcomes.

Get Clear Two: The direction of change. Some of us focus on the community, with the belief that if the community is improved, changes in the individual will occur—the community is the agent of change. Others of us believe that communities are made of individuals, and the only was to get community change is to begin with individuals—the individual is that agent of change. And then there are those who hedge their bets among us, believing that both are necessary. Both may well be necessary, so evaluators need to have clarity on which direction dominates, under which circumstances. We also need to get clear on whether we believe that services for adults actually do confer direct benefits on children. In these get clear issues, there are hidden assumptions that many of us have buried under the rug; we need to surface these, with the understanding that clarifying them will help address not only the process and direction of change, but unit of analysis issues.

Get Clear Three: Participants. All espouse the mantra of getting parents and consumers involved in programs and in evaluations. We know that such involvement can yield better data, more accurate ways of understanding the life experience of those involved in the programs, and perhaps even greater access to the participants themselves. Yet, we are loathe to verbalize the costs of such participation. Such involvement takes inordinate time; it may make the evaluation far more complex; it may distort analytic clarity; it may reduce degrees of freedom; and finally it may derail evaluators into program matters. The hard truth is that comprehensive, community-based efforts are political entities, evaluation is a political process, and the engagement of consumers is a political necessity. The question is how much engagement, when, and under what conditions? At what points is participant and consumer engagement most beneficial for the evaluation and for the consumers?

Get Clear Four: Context. We are equally enthusiastic about context. Alice O'Connor (1995) beckons us to consider context from economic, political, geographic, and temporal perspectives. She notes that a comprehensive vision of reform coupled with a focus on implementation or process demands attention to context. While conceptually accurate, such a broad-based focus on context could logjam the evaluation with polyannish promises of trying to define and take every contextual variable into consideration. Again, while considering the

vital importance of context, we need to narrow our expectations regarding what can be attributed to which contextual variables.

GETTING GOING: EXPERIMENTING WITH NEW AND PROMISING APPROACHES

Nobody in the practitioner world waited for the perfect program model to begin implementing comprehensive, community-based programs. Rather, they took the plunge and got started, using their best knowledge and understanding of the issues. They recognized the iterative process of program implementation, as evaluators need to realize the iterative nature of advances in evaluation design and methodology. Nonetheless, the lack of perfection cannot stultify work. We need to get going in several areas that respond to the issues raised above.

Get Going One: Outcomes. Getting going on outcomes is important in all evaluations, but in individually oriented interventions the treatment can be specified first with the evaluation set up to examine differences in outcomes. By contrast, community-based interventions—given the breadth of their purview and the diversity of the interventions—must specify outcomes at the outset. And only by specifying outcomes early on will program implementors be able to discern whether their varied interventions are appropriate to the desired ends. As noted earlier, however, discerning outcomes may not be quite so easy because there are at least four different kinds or types of outcomes related to these efforts.

Type One: This category includes information on what *children and families know and can do*. For children, such information must be gathered by observing them directly so that data represent a solid, precise reflection of children's performance. Behaviors in this type include dimensions related to children's motor development, their social and emotional development, their use of language, their cognition and general knowledge, and the way in which they approach learning. For adults, this type represents what adults know about various dimensions of their lives—services, parenting skills, technical skills and may be evaluated by means of observation or adult questionnaire.

Type Two: This category contains information regarding the *conditions* that surround and encase what children and families know and can do. Such information may be gathered from reviews of documents (including health records), interviews with family members and service providers, and direct observations/conversations with children and their families. Rather than reporting data on individual children or families, this type generally reports data based on aggregated prevalence and percentages. Child and family conditions may be grouped into categories—for example, child health conditions or family income conditions, with positive and negative indicators in each.

Type Three: This category contains information on the services that exist and those to which children and families have access. Distinct from the behaviors (Type One) or conditions (Type Two), this type is called the *service provision and access* type. More than a tally of raw services, this type focuses on actual access to services, with items typically reported in prevalence or percentages. Often indicators include access to services by specific

populations or individuals with particular conditions, e.g., handicapped children, pregnant women, unemployed mothers. Data for this type are typically collected from record reviews and community and institutional data bases. Examples of the information in the provision/access category include: health provision/access; parenting education provision/access; child care/preschool provision/access.

Type Four: This category contains information on the capacity of systems to perform as integrated entities. Rather than focusing on the provision of and access to discrete services as indicated above or looking at the efficacy of individual service domains, this type is far more developed than other types. It includes examinations of service redundancies, omissions, capacities, and efficiencies. Data for this type of outcome are collected in the aggregate and typically involve the amalgamation of information across agencies and service providers. Examples of categories include: systemic efficiency; systemic infrastructure; systemic accountability.

Getting going on outcomes means being precise about how much is needed to be known about each category—e.g., discerning a balance among the types. It does not mean that all comprehensive, community-based efforts will focus on the same types or even the same items within types. It does mean that each effort will consider all outcome types, and discern which information is appropriate to its goals.

Getting Going Two: The direction of change, participants, and context. Carol Weiss (1995) has advanced some very helpful work that allows evaluators to closely examine the direction and nature of change. Weiss believes that social programs are based on explicit or implicit theories about how and why programs work. These assumptions suggest that theories of change undergird program implementation and serve as a basis for examining program accomplishments. The aim of the evaluation is to discern these theories and then to use the evaluation to understand how the theories work, which of the assumptions hold true, which break down, and under what conditions. The presumption is that this approach will help clarify what exactly is to be evaluated.

Working collaboratively with parents and consumers, the theories of change approach can enlist opinions and ideas from a number of participants who gather to discuss program intentions, activities, and the link between intentions and activities. Outcome pathways and models of change unique to each program can be constructed based on information from participants. Finally, outcomes can be specified and instruments to measure such outcomes selected or constructed.

Using a theory-driven approach to understanding and specifying change can be advantageous for several reasons. First, the theories of change work can lend precision to what is to be evaluated and it can lend precision to program efforts. Second, it can give researchers and practitioners a chance to grope together, thereby establishing a new tone of reciprocity and trust. Third, the process can also accommodate variation in context while enabling it to be examined in accord with the setting. In this sense, theory-based evaluation can meet three

goals: clearly charting the direction of change; engaging participants in ways that build and sustain collaborative, productive relationships; and accommodating contextual variation. Clarifying outcomes and engaging in theories of change work are two promising "get goings." They are not the only ones. Workers in the evaluative field are engaged in much promising work that links qualitative and quantitative work, that uses time series approaches to chronicle changes over time, that compares exemplary and typical practices, that uses ethnographic techniques or data from new management information systems, that finds inventive ways to make cross-site comparisons, or that combines several of the above. Whatever evaluative strategies are used, the real point is that we need to rethink not only our evaluation practices, but the purposes for which evaluations are being conducted. It is clear that old paradigms, with their attendant purposes, are no longer sufficient.

REFERENCES

Behrman, R. E. (Ed.). (1992). School-Linked Services. Los Altos, CA: Center for the Study of the Child.

Bruner, C. (1994). A framework for measuring the potential of comprehensive services strategies. In N. Young, S. Gardner, S. Coley, L. Schorr, & C. Bruner (Eds.), Making A Difference: Moving to Outcomes-Based Accountability for Comprehensive Service Reforms, (pp. 29-40). Falls Church, VA: National Center for Service Integration.

Connell, J. P., Kubisch, A. C., Schorr, L. B., & Weiss, C. H. (Eds.). (1995). New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts. Washington, DC: The Aspen Institute.

Coulton, C. J. (1992). Framework for Evaluating Family Centers. Cleveland, OH: Case Western Reserve University, Center for Urban Poverty and Social Change.

Crowson, R. L. & Boyd, W. K. (1993). Coordinated services for children: Designing arks for storms and seas unknown. American Journal of Education, 101, 140-179.

Knapp, M. (1995). How shall we study comprehensive, collaborative services for children and families? Educational Researcher, (24)4, 5-16.

O'Connor, A. (1995). Evaluating comprehensive community initiatives: A view from history. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts, (pp. 23-64). Washington, DC: The Aspen Institute.

Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts, (pp. 65-92). Washington, DC: The Aspen Institute.

NEW APPROACHES TO EVALUATION:
Helping Sister Mary Paul, Geoff Canada and Otis Johnson
While Convincing Pat Moynihan, Newt Gingrich and the American Public[2]

Presentation by Lisbeth B. Schorr[3] and Anne C. Kubisch[4]

at the
Annie E. Casey Foundation Annual Research/Evaluation Conference
September 29, 1995

At the outset, I want to say a word about the current political context, and why I think it is at all worth talking about research and evaluation in support of promising programs and policies at a time when the structures of support that we had long considered a permanent part of our national life are being dismantled. Dismantled in an outpouring of bipartisan meanness, propelled by some bizarre illusion that it's the poor who are responsible for the trouble this country is in. Clearly the nation, always wary of activist government, has hit a new high in citizen distrust and legislative hostility toward governmental efforts, especially those meant to help the disadvantaged, who are thought—secretly and tentatively by liberals and vocally and certainly by conservatives—to be somehow responsible not only for their own misfortune, but for whatever disturbs the rest of us, be it high levels of crime, stagnating wages, or high taxes.

But this era of meanness is not going to last forever. There may soon be a revulsion against the destruction we are now living through, and we may soon have some new opportunities. The current ferment and discontent with our social institutions has vastly expanded what is discussible and—for better or worse—what is regarded as changeable.

I believe that those of us who recognize that we need a new sorting—to distinguish between the solutions and the institutions that are working and those that are not—have to be able to propose alternatives to the now popular notion that our most serious social problems will be solved by dismantling government or reducing government to a punitive force and leaving the unfettered market and private charity to cope with the problems that government has not been able to solve. I see a new urgency to our search for ways to make all our institutions work more effectively, be they public, private, or some new combinations of the two.

---

[2] This presentation is based on a chapter from a forthcoming book by Lisbeth B. Schorr, tentatively titled High Stakes: Families, Communities, and the National Future.

[3] Lecturer in Social Medicine, Harvard University; Director, Harvard Project on Effective Services.

[4] Director, Roundtable on Comprehensive Community Initiatives for Children and Families, Aspen Institute.

In this task, the people gathered here today have a major role to play because all of you are connected in some way to the most promising efforts to build communities and to reform institutions in order to improve outcomes for high risk children and youth.

More specifically, everyone here must become part of finding better ways than we have relied on in the past to learn from these efforts about what it takes to build on what works. And the learning must take place at every level—at the program level and at the policy level. The learning at the local program level must inform efforts to reach the levers of change that can only be reached by those people who can operate at the systems level. We must be clear that while the source of the data, the insights, the wisdom about how systems must change to support effective local initiatives can only be those local initiatives, the clout and capacity to bring about systems changes in financing, regulation, accountability and governance is elsewhere. It is much more likely to be at the national and state level, and in organizations designed to deal with policy issues rather than with local services, supports and community building.

But let me return to our agenda here, the specific changes needed in research and evaluation as part of a larger strategy to improve the conditions under which high risk children grow to adulthood. I would like to submit three propositions for your consideration:

First, that anyone trying to improve the conditions under which high risk children grow to adulthood, must pay close attention to the changes needed in prevailing approaches to research and evaluation.

Second, that prevailing approaches to research and evaluation must be changed in ways that will help to improve programs **while at the same time** providing skeptics with persuasive evidence of program effectiveness. (That's what I mean about helping Sister Mary Paul, Geoff Canada, Otis Johnson and their many noble colleagues while simultaneously persuading Senator Moynihan and Newt Gingrich of the effectiveness of their efforts.)

And third, that by changing our approach to evaluation, we can bring about a **sorely needed realignment**—to get us away from a stand-off between one group of people who are seen as rigorous and objective, who are willing to focus on outcomes, and who are absolutely convinced that **nothing works**, and the people on the other side who are seen as soft and subjective, who are eager to focus on process to the exclusion of results, and who believe that well-designed interventions can change lives.

Let me offer some background for these propositions:

One of the ways in which policy makers in the United States differ from those in other countries, is that they believe—or say they believe—that social science should guide their decisions about social programs and social policy.

The evaluation profession that has been spawned by that belief exerts its influence through its promise to use the scientific method to figure out what works. That promise is the basis of its considerable influence over social policy.

And this influence, in my view, has been predominantly—not totally, but predominantly—destructive. I don't want to minimize the contributions to our understanding of certain specific interventions, such as welfare-to-work programs, where evaluations have provided a much deeper understanding about the specifics of what aspects of the intervention seem to work best and for whom. But when it comes to the broad, complex, and interactive interventions, we have been less fortunate.

Because of the narrow range of interventions that can be assessed with current evaluation techniques, and the narrow range of information about impacts that current evaluation techniques are able to capture, prevailing approaches to evaluation have not provided the knowledge needed to make good judgments about a range of social programs that may hold the most promise. But current evaluation techniques have managed to systematically bias program design and policy development away from what is likely to be most effective.

I believe that the national conviction that nothing works, the pervasive sense that nothing can be done about our major social problems, owes a lot to the fact that the evaluations that most policy makers rely on overwhelmingly favor activities where single problems are addressed by single, usually simple, and highly circumscribed remedies. And that, of course, is not where the answers lie.

When Mary McGrory reported earlier this year on a Senate Finance Committee hearing on welfare reform and teenage pregnancy, she wrote: "One certainty, and only one, emerged from (yesterday's) hearing....No one has the faintest idea of what to do about unwed teenage mothers....After two hours of articulate and thoughtful testimony from a panel of four experts who had all the latest data and theories in hand, Moynihan said humbly,...'[T]his morning we have learned how little we know and how much we have failed and how much we have denied our failure.'"

Senator Moynihan's conclusion is simply wrong. (I say this in full awareness that Sen. Moynihan may have the highest IQ of any member of the U.S. Senate, and that it takes a lot of chutzpah to second-guess him.) But I believe that Sen. Moynihan was misled by relying on studies that looked at only the narrowest of interventions, because the kind of rigorous evaluations he considers scientific have been confined to the narrowest interventions.

In my view, Senator Moynihan and his colleagues have been relying on an outmoded approach to evaluation—that has had us looking for answers in all the wrong places. (Martin Gerry, assistant secretary for policy and evaluation in the Department of Health and Human Services in the Bush Administration, likes to say that the reason I was able to find all those programs that worked that I wrote about in *Within Our Reach* is that I didn't rely exclusively on the formal evaluation literature to figure out what works. He's right.)

The conventions that have governed impact evaluations have systematically defined out of contention precisely the interventions that sophisticated funders and program people consider most promising.

To be rigorously evaluated with traditional methods, interventions must be
- standardized and uniform—across sites and across families and individuals, and over time; they must be
- sufficiently circumscribed that their activities and effects can be discerned in isolation from other attempts to intervene and from changes in community circumstances; and they must be
- sufficiently susceptible to outside direction that a central authority is able to design and prescribe such features as how participants are recruited and selected.

But of course these are precisely the conditions that have been found to be incompatible with program effectiveness. Effective programs are adapted to respond to particular sites, families, and individuals; they change over time, with continuing mid-course corrections to raise the odds of success; they are comprehensive, complex, interactive, and multi-faceted; they include efforts to change community conditions; they recognize their dependence on macro-economic and other large social forces; and they count on being able to make operational decisions locally.

So how did we get into this mess, where the very characteristics that make for good programs also make them "unevaluatable?"

As long ago as 1976, Alice Rivlin—now the director of OMB—warned that "maybe the whole evaluation movement started off on a couple of false premises...that there is such a thing as a social program, in the sense of a treatment, which applied [equally] to [all] people, which can then be evaluated to see if it works or not. Most of the evaluations...assumed that we were providing something to people, that we could say what it was, that we could define some sort of output, and that we could measure whether it took place or not." There have been few challenges to this evaluation mindset in the intervening years.

In my view, the origin of the problem was the fledgling evaluation industry's reliance on the bio-medical, experimental model as the sole basis for understanding social and human service programs. This model, which is really only useful in those instances where the intervention to be tested works just like penicillin, "assumes the presence of a pre-made service, [a uniform treatment] that...need only be administered in the right dosages to ensure success for interchangeable customers. The client may—indeed should—remain patient and passive until his or her medicine arrives....What is given is presumed equivalent to what is received, and what is received is equal to what is used. Use is then equated to gain."

Once one assumes a uniform, standardized "treatment," the requirements imposed by the bio-medical model of evaluation make sense. In the context of a uniform treatment that is independent of interactions among the persons involved, random recruitment and selection of

subjects (which allow for unambiguous comparisons between those receiving the "treatment" and a statistically similar group who do not), make sense. Assuming a standardized intervention that can and should be held constant across sites and over time, makes sense. Dismissing the effects of variations in neighborhood environments as "contaminants" makes sense.

It even seemed to make sense for evaluators to tell program people that if the intervention they designed did not fit into this Procrustean bed, they should change their program design to make it "evaluatable."

So what happened to the programs with the attributes most likely to change outcomes—because they were customized to respond to individuals and families "with subjective interiors, wants, dislikes, and ambivalences," and to respond to diverse communities each with their own needs and strengths, because they consisted of many parts that interacted with one another, because they were designed to change environments and not just individuals? Such programs were either deemed not "evaluatable" and therefore not evaluated, or they were judged not to be susceptible to an impact evaluation, and therefore subject to a process evaluation only, or their sponsors were persuaded to simplify and narrow them and standardize them to make them "evaluatable"—and then, lo and behold, they were found to be unsuccessful in changing outcomes!

I believe that the big funders, public and philanthropic, whose pressures were shaping evaluation, might have questioned their assumptions and seen the folly of constraining interventions in that way, had they not been so eager to seem as hard-headed as their physical and biological scientist colleagues. But they concluded that the aura of science and the sheen of certainty that the early evaluators offered made up for any constraints the evaluators might impose on program design. They thrived on their emulation of the "hard" sciences, on using an experimental approach that would "approximate a laboratory setting as closely as possible." They built their reputation for scientific objectivity on experimental design and became "preoccupied with its requisites: finite, measurable program goals; discernible program components; the ability to control for internal and contextual contingencies; and generalizability across locality."

Soon legislation was being passed that specified that new social programs must be evaluated as a condition of continued funding, and that evaluations must use an experimental design with randomly assigned control groups. Only in this way could policy makers be confident that the observed impacts were indeed the result of a designated treatment. If that left no room for community building, for strengthened social bonds, for unique responses to unique circumstances, and the possibility that the whole of the intervention would turn out to be more than the sum of its parts, it was assumed that little would be lost. By 1988 the Urban Institute's Isabel Sawhill was writing that "a consensus seems to be emerging that...random assignment should be the *sine qua non* of future evaluations."

The teenage pregnancy prevention programs that formed the basis for Sen. Moynihan's grim conclusion were those that aimed at changing only a single element in complex adolescent lives—by increasing access to contraception, by improving education about sexuality, or by trying to convince youngsters that they would have a better future if they postponed sexual activity or childbearing. None of the interventions that the Senate Finance Committee heard about that day represented efforts to combine all three of these elements and to add a fourth: changing the circumstances of the youngsters' lives to raise the chances that they would actually have a better future.

Let me ask you to join me in a thought experiment: suppose it turns out that what many of us believe to be true about reducing teenage pregnancy is in fact true, even though we have no proof. My contention is that with the present state of affairs, we would have no way of producing the evidence.

What do we know? We know that given prevailing inducements to engage in early sexual activity, and given prevailing methods of contraception, postponing sexual activity and avoiding pregnancy is a complicated, challenging task that requires consistent dedication over an extended period of time. Even a fleeting step off the straight and narrow can result in pregnancy. We know that for young women with few alternatives and little hope, to whom a baby offers the promise of unconditional love, a chance to feel needed and valued, and a feeling of accomplishment, the calculus of choice is more complex than legislators or editorial writers like to admit. A recent report from the Institute of Medicine points out that because the human organism is designed to reproduce absent the utmost vigilance, the motivation to avoid unintended pregnancy must be extremely powerful if pregnancy is to be prevented. Those who are ambivalent about childbearing turn out to be at just as high a risk of having a child as those who positively desire to conceive. The IOM report concludes that "Hopes and plans for a better adult life—**and reason to believe that the plans are realistic**"—are what it's going to take to overcome all the many obstacles that poor young women face to remaining abstinent or using contraception successfully.

"Reason to believe that the plans are realistic" obviously means going well beyond what we usually think of as teenage pregnancy prevention.

- It means the availability of high quality schooling to make sure that disadvantaged young men and women will have the skills and motivation needed for employment;
- it means making sure that there are decent jobs out there that pay a living wage, and that the connections are there to link these young people to job opportunities—connections that don't now exist, between school and work, and between isolated ghettos and employers who are hiring; and
- it means communities that support families in their childrearing and young people as they struggle to find their way into a healthy adulthood and into a society in which they have a stake.

Now suppose that line of argument turned out to be correct. By looking to current evaluation research, we wouldn't find a clue that that might be so.

No intervention designed in accordance with those findings would have found its way into Senator Moynihan's orbit, because it would have been found too complex, too interactive, and too messy for an experimental design, and therefore "unevaluatable."

Harvard Law Professor Martha Minow came to a similar conclusion in trying to understand why there was so little "social scientific evidence" documenting the effectiveness of home visiting programs to families with infants. Her review of these programs led her to believe that they had been highly successful: they provided support at times of stress, improved the health status of the children, and increased the economic independence and self-reliance of the parents. But social science findings were not providing policy makers with the kind of evidence they needed to scale up public support for home visiting. She concluded that "the very cautiousness of social science undermines its usefulness in policy making" by limiting what counts as reliable knowledge and rejecting as untrustworthy studies that fail to use randomized assignment.

And of course cautiousness and skepticism is what gets you respect not just in the academy, but also among legislators. So we have Peter Rossi, famous among evaluators in part because as long ago as 1978, he promulgated as Rossi's Iron Law that "The expected value for any measured effect of a social program is zero." The same Peter Rossi who came to the first meeting of the evaluation steering committee of the Roundtable on Comprehensive Community Initiatives and announced, in response to my description of the mismatch between the most promising interventions and prevailing evaluation approaches, that if such a mismatch did indeed exist, then program design would have to change!

The fundamental mismatch between prevailing evaluation approaches and the most promising kinds of interventions has resulted in

- a skewing of program design away from complex, interactive, responsive, evolving, community-based interventions—in the interest of making the intervention "evaluatable," and
- a lack of reliable information about many interventions that have in fact been successful, but that have been considered "unevaluatable."

The sparsity of information that would allow the public and policy makers to judge the effectiveness of major new social policies and programs, and that would allow programs, communities, and policy makers to understand, learn from, and improve complex cross-systems interventions is a direct consequence of the evaluation conventions that have been venerated for too long, despite the fact that they have not served the nation well.

But the near unanimous acceptance of prevailing evaluation approaches may be coming to an end. As more and more researchers, practitioners and funders came to appreciate the

importance of neighborhoods and of community engagement, of the interaction among interventions, and of elements that are difficult to quantify, it became glaringly apparent that any method of evaluation that excludes such factors from its domain could not long be considered legitimate.

After all, if Robert Putnam's now famous *Bowling Alone* analysis, incorporating research findings like those of Larry Katz and Anne Case, can show that an adolescent's chance of being arrested decreased *if he had a neighbor who was a churchgoer*, the penicillin analogy mindset has to go.

A lot of tough-minded people are gradually becoming more open to reexamining earlier conceptions of interventions as standardized treatments administered to one individual at a time, and are coming to wonder whether they've been paying too high a price in trading off the opportunity to obtain a rich array of policy-relevant information against methodological elegance and certainty.

It is no accident that one of the first national efforts to try to develop alternative evaluation approaches is being undertaken by a group created to understand and learn from the experiences of comprehensive community-based initiatives, the Roundtable on Comprehensive Community Initiatives for Children and Families. A report on the first phase of that work concluded that funders should continue to press for evidence that the interventions they are supporting are accomplishing the objectives for which they have been funded, while being mindful of the fact that significant change takes a long time, and that their standards of certainty of evaluation information may need to be revised. The Roundtable recommended that new approaches to evaluation be developed, since so many individual interventions are necessary but not sufficient to improve outcomes. If the most promising efforts are made up of several initiatives, operating in the same community under separate auspices, their combined impact—no matter how significant—simply cannot be judged with prevailing approaches to evaluation.

But it will not be easy to move away from the conventional dogma. After all, experimental designs using random assignment are the way to be most certain that the intervention being tested is what caused the difference in outcomes between the participants in an experiment and the control group. That is why Swarthmore Economics Professor Robinson G. Hollister, a leading figure in evaluation circles since the mid-1960s, says that experimental designs are "a bit like the nectar of the gods: once you've had a taste of the pure stuff it is hard to settle for the flawed alternatives."

The alternatives, however, while flawed because they provide less certainty, have the potential of providing a great deal more useful information about what really matters. The alternatives are based on the assumption that there is knowledge that is worth having and acting on even if it is not absolutely certain knowledge. They assume that policy making requires knowledge that includes, and does not reject on grounds of messiness, information

that can shed light on the complexities of human connections, and on real-world interactions among individuals, families, and communities.

The most promising alternative approach to traditional evaluation would use a combination of several traditional tools, but supplement and strengthen those tools by grounding them in the theories of change that underlie the initiatives that appear to have the greatest likelihood of success. Theory-based evaluation posits that where statistical analysis alone cannot provide the needed answers about what's working, numerical measures of outcomes combined with an understanding of the process that produced the outcomes, can shed light simultaneously on the extent of impact and on how the change occurred. Returning once more to my theme, theory-based evaluation can help to persuade the skeptical funder, legislator and taxpayer, while helping the people in the frontlines to improve their programs and allowing others to learn from their successes.

Although the ideas underlying theory-based evaluation have been around for some time, it has only recently been applied to the evaluation of complex interventions. Huey-Tsyh Chen and others have applied theory-based evaluation in the substance abuse arena, and more recently, Professor Carol Weiss, of the Harvard Graduate School of Education, has been working on a further development of theory-based evaluation and its application to comprehensive community initiatives. She explains that because all interventions are based on theories—which may be implicitly or explicitly held—an essential beginning to understanding is to identify the operative theory or theories about the things in a program or initiative that matter.

In using theory as a starting point, Professor Weiss' work is in the finest tradition of social science. The activist/academic John Gardner points out that "what is most striking about the enormously useful work of people like Darwin and deTocqueville, is that they came to their observations with very well-developed concepts. [They got away from the] fruitless efforts to measure precisely the variables which were not relevant or to answer questions which did not reflect a theory of change....They knew what they were looking for."

We too must have the courage to say that we know a lot about what we are looking for. The "well-developed concepts," or theories of change that we need can be and are currently being identified by evaluators working closely with practitioners and researchers. For example, the theory behind an effort to improve services and supports for preschool children and their families might be articulated as *children whose experiences during infancy and early childhood equip them to enter school "ready to learn" are more likely to succeed at school than children who enter school not "ready to learn" because of early deficits in health care, nutrition, child care and preschool experiences, because they lived in communities that did not support families in ways that were conducive to developing trust, curiosity, self-regulation, the foundations of literacy and numeracy, and social competence.*

To take another example, the theories underlying an effort to build up Little Leagues and youth groups and to create a community school in a previously devastated neighborhood

might include the following: *Inner city youth are more likely to finish school, have a job, and avoid drugs and crime if they have more social capital to draw on because they live in neighborhoods with high levels of civic engagement, which can be brought about as a by-product of other social activities which can be systematically encouraged and supported.*

Once the theories have been identified, the evaluator works with the program people to identify the microsteps that are hypothesized, on the basis of experience and research, to link the various parts of the theory to one another. In the case of the "ready to learn" theory, these micro-steps might include markers of community capacity such as the availability to all low income families of accessible, responsive, high quality health care for infants, children and pregnant women, child care that combines developmentally appropriate care and education and family support, child protective services, family support programs, adequate nutrition, adequate income, and a supportive community infrastructure. Other links might be identified in the form of interim outcome measures, including higher rates of pregnant women receiving prompt and continuing prenatal care; higher rates of infants and preschool children receiving preventive health care, including immunizations; higher rates of 3- and 4-year-olds in Head Start and other high quality child care/education settings; higher rates of infants and toddlers (whose families want or need out-of-home care for them) being cared for in high quality child care settings; fewer confirmed and repeat instances of child abuse and neglect; and lower rates of inappropriate out-of-home placements.

In the case of the "social capital" theory, the microsteps might include such short-term markers of community capacity as an increase in the number of community clubs and associations; attendance and participation rates; attendance at religious services; registration and voting; number of books, tapes, etc., borrowed from local libraries; and children, youth and parents using neighborhood playgrounds and other recreational facilities. Interim outcome measures might include improved school attendance, dropout and graduation rates, and performance on achievement tests; and reductions in crime, auto theft, arrests of minors, other crime statistics, and in rates of youth idle on the streets.

The theories of change approach to evaluation, then, has evaluators, practitioners, and researchers working together to construct a "conceptual map" that links all the important parts of an intervention to one another. Increasingly there will be more indicators and measurements along the entire causal chain to help participants, program people, funders and policy makers to arrive at an ever richer understanding of what is being accomplished and how it is being accomplished.

Progress along these lines will require a great deal of new work on the interim milestones that link interventions with ultimate outcomes, that could reliably show that reform efforts are on track toward achieving their targets.

After all, the most frequently cited lesson from major current reform efforts is that they take so much more time than expected—both to get the initiative under way, and to get it to the point where it begins to show an impact on real-world outcomes. We desperately need new

tools that would allow initiatives to demonstrate their short-term achievements over time periods that are meaningful to politicians, to those funders who do not easily take a long-range view, and to community residents who are eager for documented evidence of progress. They need to be able to get interim information very quickly—often long before a program is "proud," long before it has had a chance to make an impact on rates of school readiness, child abuse, teenage pregnancy, violence, school success, and employment.

Two kinds of interim measures can predict later outcomes: indicators that attach to children, families, and communities and that are a short-term manifestation of long-term outcomes, and indicators of a community's capacity to achieve the identified long-term outcomes.

Knowledge about the connections between measurable indicators of community capacity and long-term outcomes is at a more primitive stage than knowledge about the connections between interim and long-term indicators for children and families. Reliable theories about the linkages between interventions and results, and about the constellation of conditions and interventions that will lead to good results, are scarce. Most are unproven. For example, can a community that is developing strategies to reduce rates of low-weight births assume with confidence that the "enabling conditions" to reach that outcome are some combination of the capacity (1) to provide family planning services to all persons of childbearing age, and (2) to provide high quality, responsive prenatal care, nutrition services, and family support to pregnant women? Are measures of the extent of program participation, client satisfaction, or an increased sense of community reliable precursors of improved outcomes?

The availability of family planning and prenatal care and health insurance are surely related to improved birth outcomes, but whether the relationship is strong enough, and whether their effect on outcomes is actually a function of their availability (rather than of their quality), so that their **availability** can be used as an interim indicator, is an open question. It is probably not enough to know of the simple existence of certain services, because their quality and how they are made available must be taken into account to link them strongly with outcomes. The distinction among service **availability, access,** and the **nature and quality of the service** in accounting for improved results is crucial—and requires greater understanding and a wider consensus around how to measure the factors that make services effective than now exist.

Perhaps the most tantalizing of recently hypothesized links between interventions and outcomes that could produce some new short-term indicators of community capacity are between outcomes for children and families and such indicators of community-level change as a strengthened infrastructure of informal supports, and investments in neighborhood safety and expanded economic opportunity. But there is as yet scant agreement on ways to measure community building, and only modest understanding of the precise connections.

The need for both kinds of short-term indicators that could show movement toward long-term outcomes has long been recognized. It has not been met because the ability to define these interim markers with confidence depends on having reliable evidence, theories, or at least

sturdy hypotheses, about the antecedents of major long-term outcomes. Neither social science researchers nor the evaluation industry have really invested in this arena—in part because progress in this arena involves a higher ratio of judgment to certainty than most social scientists are comfortable with.

As a society, we now need desperately to make up for lost time. One useful next step would be to systematically examine findings in the recent literature and ongoing experience to provide a more rigorous and deeper understanding of established connections among short-term and long-term outcomes. We need to explore the connections between long-term outcomes on the one hand, and measures of interim individual outcomes and community capacity on the other. The evaluation steering committee of the Aspen Roundtable on Comprehensive Community Initiatives has been discussing the usefulness of a "Michelin Guide" to interim indicators, that would assess the degree of confidence with which the hypothesized connection between interim indicators and long-term outcome measures could be linked, all along the causal chain. The idea would be to distinguish among the connections that seem to be fairly well established, those where the evidence is weaker and the hypothesized connections urgently need to be tested, and those where even promising hypotheses are lacking.

Although these are the basic components of a theories of change approach, evaluators are going to have to supplement them with some of the conventional tools of evaluation (including comparisons among populations and communities, comparisons over time, etc.) where those can be applied without distorting program design.

All of the approaches relying on quantitative data must also be linked to the work of ethnographers and other sophisticated observers who will document and describe the successes, failures, and processes through narrative. The detailed and subtle narrative, as Professor Sara Lawrence-Lightfoot makes clear in her own inspiring work and in her challenge to her colleagues, can be the "thick description" that "allows us to see the interaction of the key ingredients of change, and to record the experiences of those who are engaged in the process." The narrative allows us to see into the relationships that are at the core of good practice.

Funders and program people should not have to choose between achieving a greater understanding of process or impact. Both are essential. The problem arises when the information about process is used as a **substitute** for information about impact. This is the phenomenon that David Osborne calls "process creep."

When process creep occurs, means and ends become confused, and the focus on what actually happens to people as a result of the activity is lost. The formation of a collaborative, or a high degree of participation in a new governance entity may be the product of a great deal of effort, but is not evidence of progress toward agreed upon outcomes unless the rationale that connects these activities to established outcomes is at least explicitly hypothesized, if not proven. The number of children who have been screened for hearing and vision problems is a process indicator. Because screening that isn't followed up with diagnosis and treatment where needed won't reduce the number of children whose vision or hearing is impaired, screening

should not be used as an outcome indicator. The vocational school that pumps out ever larger numbers of certified welders, even though the school's graduates cannot find jobs because robots have replaced welders, is not achieving valued outcomes.

But the critical confusion about process measures is not conceptual, it is political. The temptation is ever-present to fall back on using process measures as evidence of progress, even when they meet none of the criteria for outcome measures and there is no basis for linking them to ultimate outcomes. Process measures so often become substitutes for outcome measures because they provide comforting evidence of activity, they demonstrate that **something** is happening.

Typically, both grantmakers and grantees contribute to process creep. It happens in the early stages of program implementation, when everyone involved suddenly becomes afraid that his or her hopes for the project may not be realized, and begins to view evaluation research as an "unfriendly act." It also happens when funders encounter hostility to outcome accountability (and outcome evaluation) from communities and program people who fear that outcome measurement will not do justice to their underfunded intervention.

In responding to these fears, funders often find it easier to move or remove the goal posts than to strengthen the players.

The typical forget-about-the-goal-posts conversation takes place a few months into the implementation phase of a program. The funder says to the grantee something along the following lines: So we gave you the grant in the hope that you would reduce teenage pregnancy and youth violence in this community, and you now say that was really an unrealistic expectation? You may be right. But we do need some hard evidence that our grant is making some sort of difference, so let's see if we can get an evaluator to design an attitude survey that will determine whether you have increased the number of teenagers who think it's a bad idea to carry a gun and to initiate sex when they're younger than fifteen. Or the evaluators could document how many youngsters come to your meetings and classes. Alternatively, maybe we or you could hire an ethnographer to chronicle what's going on in your program....

Some of these are useful things to do. It is especially useful to obtain rich descriptions of complex, nuanced interventions. But descriptions of process are most useful to program people as well as to funders and policy makers when they become an integral part of a rigorous and systematic inquiry into what the program is accomplishing and why.

A greater focus on outcomes and results may have its most profound effect by calling attention to whether investments are adequate to achieve the projected results. An outcomes focus injects what Sid Gardner calls a strengthened ethical core into human service systems that currently focus more attention on the fate of agencies and programs than on whether people are actually being helped. The new outcomes focus promises (or threatens, in the eyes of some) to end a conspiracy of silence between funders and program people by exposing the sham in which human service providers, educators, and community organizations are consistently asked to

accomplish massive tasks with inadequate resources and inadequate tools. Attention to results forces the question of whether outcome expectations must be scaled down, or interventions and investments scaled up to achieve their intended purpose.

In the past, parent education programs have been funded in the vague expectation that they would somehow reduce the incidence of child abuse, although a few didactic classes have never been shown to change parenting practices among parents at risk of child abuse. Similarly, outreach programs to get pregnant women into prenatal care are expected to reduce the incidence of low-birth weight in the similarly vague belief that outreach programs are a good thing without any reference to whether the prenatal care which is made more accessible actually provides the services that could be expected to result in a greater number of healthy births.

Especially in circumstances where it will take a critical mass of high quality, comprehensive, intensive, interactive interventions to change outcomes, where effective interventions must be able to impact even widespread despair, hopelessness and social isolation, funders and program people should resist the temptation to obscure the limitations of so many current efforts. Providers—and even reformers—who are asked to achieve grand outcomes with interventions so paltry that they are in no way commensurate to the task, should not be obscuring the insufficiency of the investment by pleading with funders and evaluators to just document their efforts and not their results because it wouldn't be fair to hold them accountable for real outcomes changes when they're doing the best they can. Evidence that a diluted form of a previously successful intervention is not making an impact is not an argument against results-based accountability. It helps to clarify that dilution regularly transforms effective model efforts into ineffective replications. Recognition that a single circumscribed intervention may not be sufficient to change outcomes is not an argument against results-based accountability. It is an argument for adequate funding of a combined critical mass of promising interventions.

In focusing on impacts and combining new and old approaches to evaluation, the new evaluators may offer less certainty—especially about causal attribution. But the information they bring to the table can be not only rich but also rigorous. And that rich and rigorous data can provide a solid basis for insight and further learning, and thereby lead to effective action on urgent social problems.

There can be no **scientific certainty** about remedies for youth violence or alienation, for family dissolution, for school failure, for substance abuse, or for growing childhood poverty. But there **can** be systematic and cumulative learning. Progress toward solving these problems will come through the thoughtful, structured collection and analysis of information and experience that will lead to ever greater understanding of all the promising ways to intervene. Carefully crafted, well-informed, and thoughtful approximations about what seems to be working will provide better signposts and more usable knowledge than elegant statistical analyses of trivia. And when it's done right, those signposts and that knowledge should be equally useful to the people trying to design and operate interventions and the people who are making policy and allocating resources.

In concluding, I would address those who still harbor grave doubts and a visceral unease about the whole idea of outcomes accountability and impact evaluation. Committed practitioners have every reason to ask, "Why should we have to prove the value of our work?" They point out that those who would dismantle the safety net and the whole infrastructure of public and nonprofit services and institutions aren't arguing efficacy—they're arguing principle. Many practitioners, along with parents, community leaders and other advocates, wish to stand their ground on principle, and say that feeding young children and providing them with a safe and happy place to play is enough justification, that comforting a frightened adolescent needs no further rationale, that every expectant mother is entitled to the highest quality prenatal care—regardless of whether there is a payoff in higher rates of school readiness, employability, or healthy births. Other countries, after all, don't make public support for basic services for children and families contingent on proof of their merit. In France and Germany and Britain and Japan, publicly supported child care and maternal and child health care, paid family leaves, and universal child protective services are taken for granted and require no evidence of effectiveness.

American human service leaders see themselves as part of a tradition of service to the vulnerable whose value is ultimately independent of its effects. They cite Mother Theresa's explanation of her perseverance in the face of the enormity of world poverty: "God has called on me not to be successful, but to be faithful." They cite Ghandi's teaching that "It is the action, not the fruit of the action, that is important."

My own belief is that the moral underpinnings for social action, especially by government, are not powerful enough today, in the mean and cynical closing years of the twentieth century, to sustain what needs to be done on the scale at which it needs to be done. In this time of pervasive doubt, we have to be able to provide hard evidence that investments are achieving their purpose and contributing to long-term goals that are widely shared, if we are to have any hope of obtaining the magnitude of public investment that is required.

We have to make sure that analytic rigor, objectivity, and an outcomes focus are not monopolized by the people who believe that nothing works. We have to make sure that when the current spasm of meanness relinquishes its hold on the body politic, the people who believe that well-designed interventions can indeed change lives will produce the rigorous, usable knowledge which will become the foundation for large-scale support of the interventions that will succeed in restoring hope to the children and families that now have no stake in the American dream.

Appendix B:
Conference Agenda

September 27

5:30-6:30 p.m.	Reception
6:30-8:30 p.m.	Dinner/Welcome Speech Tony Cipollone

September 28

8-8:30 a.m.	Continental Breakfast
8:30-8:45 a.m.	Charge of the Meeting and Plan for the Day Michael Grady
8:45-9:30 a.m.	Plenary Session: "Evaluating Comprehensive Community-Based Services: Rethinking Purpose and Practice" Sharon Lynn Kagan
9:30-11 a.m.	Program Panel Manuel Gutierrez Role of research and evaluation in program development and operations: How funders, administrators, and researchers can collaborate in the use of data and findings to design, manage, and improve programs. The panel is composed of researchers and program administrators involved in collaborative projects, including Tom Dewar of Rainbow Research, Dorothy Coleman of the Center for Successful Child Development, Gary Walker of P/PV, Sandy Weinbaum of AED, and Heather Weiss of the Harvard Family Research Project.
11-11:15 a.m.	Break
11:15 a.m.-12:45 p.m.	Breakout Sessions: 1. Issues in the evaluation of school-linked service programs (Joy Dryfoos) 2. Issues in the evaluation of human service system reform (Charlie Bruner) 3. Ethnographic research in urban minority communities (Susan Greenbaum) 4. Issues in the evaluation of institutional change in community development initiatives (Anne Kubisch)
12:45-1:30 p.m.	Lunch
1:30-3 p.m.	Poster Session on AECF-Funded Research and Evaluation

3-4:30 p.m.	Breakout Sessions

5. Assessing the community context and outcomes of comprehensive community change initiatives (Claudia Coulton)
6. Conducting survey research in and with empowered communities (Mary Achatz)
7. Issues in the evaluation of family support and preservation programs (Fran Jacobs)

September 29

8-8:45 a.m.	Continental Breakfast
8:45-9 a.m.	Plan for the Day
9-10:45 a.m.	Policy Panel Cindy Guy

Role of research and evaluation in public policy: How to create a compelling case for policy reform, using imperfect methods and findings.

The panel will include representatives of federal, state, and local government including Don Crary of Little Rock New Futures, David Ellwood of the JFK School, Penny Sanders of the Kentucky Office of Educational Accountability, and Ralph Smith of the Annie E. Casey Foundation.

10:45-11 a.m.	Break
11 a.m.-12:30 p.m.	Breakout Sessions

1. Ethnographic research in urban minority communities (Susan Greenbaum)
2. Experimental and quasi-experimental approaches to the evaluation of comprehensive child and family service programs (Lynn Usher)
3. Conducting survey research in and with empowered communities (Kimberly Huff)

12:30-1:30 p.m.	Lunch
1:30-2:15 p.m.	Plenary Session: "New Approaches to Evaluation: Helping Sister Mary Paul, Geoff Canada, and Otis Johnson While Convincing Pat Moynihan, Newt Gingrich, and the American Public" Lisbeth Schorr (delivered by Anne Kubisch)
2:15-2:30 p.m.	Closing Remarks Tony Cipollone

Appendix C:
Conference Participants

Mary Achatz
Public/Private Ventures, Inc.
2005 Market St., Suite 900
Philadelphia, PA 19103
(215) 557-4400

Anita Baker
Academy for Educational Development
100 Fifth Ave.
New York, NY 10011
(212) 243-1110

Nannette Banks
National Council on Crime and Delinquency
685 Market St., Suite 620
San Francisco, CA 94105
(415) 896-6223

Ira Barbell
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

John Bentley
Dept. of Social Services
2650 California St.
Chicago, IL 60608
(312) 890-6010

Cheryl Boykins
Center for Black Women's Wellness
177 Windsor St., SW
Room 309
Atlanta, GA 30312
(404) 688-9202

Pat Brennan
New York City Criminal Justice Agency
52 Duane St., Third Floor
New York, NY 10007
(212) 577-0519

William Brown
Florida Mental Health Institute
University of South Florida
13301 Bruce B. Downs Blvd.
Tampa, FL 33612-3899
(813) 974-2897

Charles Bruner
Child & Family Policy Center
Fleming Building, Suite 1021
218 Sixth Ave.
Des Moines, IA 50309
(515) 280-9027

Thomas Burns
OMG Center for Collaborative Learning
1528 Walnut St., Suite 805
Philadelphia, PA 19102
(215) 732-2200

Deborah Busch
Metis Associates, Inc.
80 Broad St., Suite 1600
New York, NY 10004-2209
(212) 425-8833

Linda Carmona-Sanchez
The East Little Havana/Casey
Children's Mental Health
1401 SW First St., Suite 111
Miami, FL 33135
(305) 649-6449

Tony Cipollone
Annie E. Casey Foundation
Research and Evaluation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Kristin Coffey
Annie E. Casey Foundation
Research and Evaluation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Dorothy Coleman
Center for Successful Child Development
4848 South State
Chicago, IL 60609
(312) 373-8670

Claudia Coulton
Center for Urban Poverty and Social Change
Case Western University
10900 Euclid Ave.
Cleveland, OH 44106-7164
(216) 368-2304

Don Crary
New Futures for Little Rock Youth
209 West Capitol Ave., Second Floor Annex
Little Rock, AR 72201
(501) 374-1011

Michael Dalmat
Centers for Disease Control and Prevention
4770 Buford Highway NE (K-20)
Atlanta, GA 30341-3724
(404) 488-5136

Deborah Daro
National Committee to Prevent Child Abuse
332 South Michigan Ave., Suite 1600
Chicago, IL 60604-4357
(312) 663-3520

Debra Delgado
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Tom Dewar
Rainbow Research
621 West Lake St.
Minneapolis, MN 55408
(612) 824-0724

Joy Dryfoos
Independent Consultant
20 Circle Dr.
Hastings-on-Hudson, NY 10706
(914) 478-3489

David Ellwood
Harvard University
John F. Kennedy School of Government
Cambridge, MA 02138
(617) 495-1121

Michael Epstein
Dept. of Educational Psychology
Northern Illinois University, Graham Hall
DeKalb, IL 60115
(815) 753-8443

Kathleen Feely
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Leila Fiester
Policy Studies Associates
1718 Connecticut Ave. NW, Suite 400
Washington, DC 20009
(202) 939-5339

Robert Friedman
Florida Mental Health Institute
University of South Florida
13301 Bruce B. Downs Blvd.
Tampa, FL 33612-3899
(813) 974-2897

Michelle Gambone
Public/Private Ventures, Inc.
2005 Market St., Suite 900
Philadelphia, PA 19103
(215) 557-4400

Deborah Gibbs
Research Triangle Institute
P.O. Box 12194
Research Triangle Park, NC 27790-2194
(919) 541-6357

Miriam Gonzales
Policy Studies Associates
1718 Connecticut Ave. NW, Suite 400
Washington, DC 20009
(202) 939-5338

Michael Grady
Annie E. Casey Foundation
Research and Evaluation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Susan Greenbaum
Dept. of Anthropology
University of South Florida
Tampa, FL 33620
(813) 974-0777

Lynn Gregory
Academy for Educational Development
100 Fifth Ave., Second Floor
New York, NY 10011
(212) 243-1110

Manuel Gutierrez
Annie E. Casey Foundation
Research and Evaluation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Marcella Gutierrez-Mayka
Florida Mental Health Institute
University of South Florida
13301 Bruce B. Downs Blvd.
Tampa, FL 33612-3899
(813) 974-4657

Cynthia Guy
Annie E. Casey Foundation
Research and Evaluation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Tony Hall
Metis Associates, Inc.
80 Broad St., Suite 1600
New York, NY 10004-2209
(212) 425-8833

Ron Hardy
Kaleidoscope Satellite Family Outreach Program
1279 North Milwaukee, Suite 250
Chicago, IL 60622
(312) 278-7200

Darlene Hasselbring
National Urban League
500 East 62nd St.
New York, NY 10021
(212) 310-9018

Mario Hernandez
Florida Mental Health Institute
University of South Florida
13301 Bruce B. Downs Blvd.
Tampa, FL 33612-3899
(813) 974-2897

Peg Hess
Columbia University School of Social Work
622 West 113th St.
New York, NY 10025
(212) 854-5186

Kimberly Huff
Metis Associates, Inc.
80 Broad St., Suite 1600
New York, NY 10004-2209
(212) 425-8833

Fran Jacobs
Dept. of Child Study
Tufts University
105 College Ave.
Medford, MA 01255
(617) 627-3355

Lisa Jones
Center for Child & Family Policy
Muskie Institute
1 Post Office Square, P.O. Box 15010
Portland, ME 04112
(207) 780-5852

Sharon Lynn Kagan
The Bush Center in Child Development and
Social Policy
310 Prospect St.
New Haven, CT 06511
(203) 432-9931

Roberta Knowlton
New Jersey Dept. of Human Services
CN700
222 South Warren St.
Trenton, NJ 08625-0700
(609) 292-7901

Lauren Kotloff
Public/Private Ventures, Inc.
2005 Market St., Suite 900
Philadelphia, PA 19103
(215) 557-4400

Anne Kubisch
Aspen Institute
345 East 46th St., Suite 700
New York, NY 10017-3562
(212) 697-1226

Michael Laracy
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Irv Lerner
Metis Associates, Inc.
80 Broad St., Suite 1600
New York, NY 10004-2209
(212) 425-8833

Elena Lopez
Harvard Family Research Project
Harvard Graduate School of Education
Longfellow Hall, Appian Way
38 Concord Ave.
Cambridge, MA 02138
(617) 495-9108

May Louie
Dudley Street Neighborhood Initiative
513 Dudley St.
Roxbury, MA 02119
(617) 442-9670

Bart Lubow
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Virginia Martinez
NEWSED
1029 Santa Fe Dr.
Denver, CO 80204
(305) 534-8342

Karin Martinson
ASPE, Dept. of Health and Human Services
Room 450-G, 200 Independence Ave. SW
Washington, DC 20201
(202) 690-5653

Sr. Mary Paul
Center for Family Life in Sunset Park
345 43rd St.
Brooklyn, NY 11232
(718) 788-3500

Sr. Mary Geraldine
Center for Family Life in Sunset Park
345 43rd St.
Brooklyn, NY 11232
(718) 788-3500

Patrick McCarthy
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Brenda McGowan
Columbia University School of Social Work
622 West 113th St.
New York, NY 10025
(212) 854-5186

Harriet Meyer
Ounce of Prevention Fund
188 West Randolph, Suite 2200
Chicago, IL 60601
(312) 853-6080

Maria Monteagudo
District 11 Management Services
1605 Southwest 107th Ave., Suite 202-C
Miami, FL 33165
(305) 377-5205

Diane Moore
Center for Black Women's Wellness
177 Windsor St., SW
Room 309
Atlanta, GA 30312
(404) 688-9202

Gail Myers
Public/Private Ventures, Inc.
6666 M Peachtree Industrial Blvd.
Doraville, GA 30360
(404) 446-9617

Sara Nathanson
Policy Studies Associates
1718 Connecticut Ave. NW, Suite 400
Washington, DC 20009
(202) 939-5336

Ellen Pagliaro
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Richard Pargament
Metis Associates, Inc.
80 Broad St., Suite 1600
New York, NY 10004-2209
(212) 425-8833

Patricia Patrizi
Pew Charitable Trusts
One Commerce Square, Suite 1700
Philadelphia, PA 19103-7017
(215) 575-9050

Tom Randall
Center for Family Life in Sunset Park
345 43rd St.
Brooklyn, NY 11232
(718) 788-3500

Kay Reiss
New Jersey Dept. of Human Services
CN700
222 South Warren St.
Trenton, NJ 08625-0700
(609) 292-1617

Larry Rouse
Texas Dept. of Mental Health and
Mental Retardation
Research and Evaluation
P.O. Box 12668
Austin, TX 78711
(512) 206-4638

Penny Sanders
Office of Educational Accountability
Capitol Annex
Frankfort, KY 40601
(502) 564-8167

Donna Schmidt
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Stanley Schneider
Metis Associates, Inc.
80 Broad St., Suite 1600
New York, NY 10004-2209
(212) 425-8833

Jennifer Soderlund-Sharma
Dept. of Educational Psychology
Northern Illinois University, Graham Hall
DeKalb, IL 60115
(815) 753-8443

Gerri Spilka
OMG Center for Collaborative Learning
1528 Walnut St., Suite 805
Philadelphia, PA 19102
(215) 563-2201

Donna Stark
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Michael Sullivan
Ounce of Prevention Fund
188 West Randolph, Suite 2200
Chicago, IL 60601
(312) 853-6080

Ed Tetelman
New Jersey Dept. of Human Services
CN700
222 South Warren St.
Trenton, NJ 08625-0700
(609) 292-1617

Lynn Usher
University of North Carolina
School of Social Work
Chapel Hill, NC 27599
(919) 962-0890

Rafael Valdivieso
Academy for Educational Development
1255 23rd St. NW
Washington, DC 20037
(202) 884-8727

Dennis Wagner
National Council on Crime & Delinquency
685 Market St., Suite 620
San Francisco, CA 94105
(415) 896-6223

Gary Walker
Public/Private Ventures, Inc.
2005 Market St., Suite 900
Philadelphia, PA 19103
(215) 557-4400

Constancia Warren
Academy for Educational Development
100 Fifth Ave., Second Floor
New York, NY 10011
(212) 243-1110

Sandy Weinbaum
Academy for Educational Development
100 Fifth Ave., Second Floor
New York, NY 10011
(212) 243-1110

Heather Weiss
Harvard Family Research Project
Harvard Graduate School of Education
Longfellow Hall, Appian Way
Cambridge, MA 02138
(617) 495-9108

Lynne White
Annie E. Casey Foundation
701 St. Paul St.
Baltimore, MD 21202
(410) 547-6600

Judith Wildfire
Research Triangle Institute
P.O. Box 12194
Research Triangle Park, NC 27790-2194
(919) 541-6357

LeeAnn Williams
National Committee to Prevent Child Abuse
332 South Michigan Ave., Suite 1600
Chicago, IL 60604-4357
(312) 663-3520

Ed Wilson
OMG Center for Collaborative Learning
1528 Walnut St., Suite 805
Philadelphia, PA 19102
(215) 563-2201

Laura Winterfield
New York City Criminal Justice Agency
52 Duane St., Third Floor
New York, NY 10007
(212) 577-0519

The Annie E. Casey Foundation

BEST COPY AVAILABLE